(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
11 January 2001 (11.01.2001)

PCT

(10) International Publication Number
WO 01/02436 A1

(51) International Patent Classification[7]: C07K 14/72, C12N 15/12, C07H 21/04

(21) International Application Number: PCT/AU00/00799

(22) International Filing Date: 30 June 2000 (30.06.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
09/346,470 1 July 1999 (01.07.1999) US

(71) Applicant *(for all designated States except US)*: **COMMONWEALTH SCIENTIFIC AND INDUSTRIAL RESEARCH ORGANISATION** [AU/AU]; Limestone Avenue, Campbell, ACT 2601 (AU).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **HILL, Ronald, Johnston** [AU/AU]; Altona Avenue, Forestville, NSW 2087 (AU). **HANNAN, Garry, Noel** [AU/AU]; 32 Earl Street, Hunters Hill, NSW 2110 (AU).

(74) Agents: **OLIVE, Mark, R.** et al.; Davies Collison Cave, 1 Little Collins Street, Melbourne, VIC 3000 (AU).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *With international search report.*
— *With (an) indication(s) in relation to deposited biological material furnished under Rule 13bis separately from the description.*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: NOVEL GENETIC SEQUENCES ENCODING STEROID AND JUVENILE HORMONE RECEPTOR POLYPEPTIDES AND USES THEREFOR

(57) Abstract: The present invention provides isolated nucleic acid molecules encoding polypeptides comprising functional steroid hormone and juvenile hormone receptors, in particular isolated nucleic acid molecules which encode polypeptides comprising the *Lucilia cuprina, Myzus persicae*, and *Bemisia tabaci* ecdysone receptors and juvenile hormone receptors. The present invention further provides functional recombinant steroid and juvenile hormone receptors and recombinant polypeptide subunits thereof and derivatives and analogues thereof. The present invention further provides screening systems and methods of identifying insecticidally-active agents which are capable of agonising or antagonising insect receptor function, or alternatively or in addition, which modify the affinity of said receptors for their cellular stimuli (e.g. insect steroids or juvenile hormones) or analogues thereof, or alternatively or in addition, which act as insecticides by virtue of their ability to agonise or antagonise the activity of insect hormones.

WO 01/02436 A1

NOVEL GENETIC SEQUENCES ENCODING STEROID AND JUVENILE HORMONE RECEPTOR POLYPEPTIDES AND USES THEREFOR

FIELD OF THE INVENTION

The present invention relates generally to novel genetic sequences encoding receptor polypeptides and insecticidal modalities therefor, which insecticidal modalities are based upon non-polypeptide insect hormones and their receptors. More specifically, the present invention provides isolated nucleic acid molecules encoding polypeptides comprising functional steroid hormone and juvenile hormone receptors, in particular isolated nucleic acid molecules which encode polypeptides comprising the *Lucilia cuprina* (sheep blowfly), *Myzus persicae* (aphid) and *Bemesia tabaci* (Silverleaf whitefly) ecdysone receptors and juvenile hormone receptors. In a particularly preferred embodiment, the present invention relates to isolated nucleic acid molecules which encode the *L. cuprina, M. persicae,* and *B. tabaci* EcR polypeptide subunits or fragments thereof, or which encode the EcR partner protein (USP polypeptide) subunits of *L. cuprina, M. persicae,* and *B. tabaci.* The EcR and USP polypeptides disclosed herein associate to form functional heterodimeric ecdysone receptors or receptor analogues. The present invention further provides the *L. cuprina, M. persicae,* and *B. tabaci* EcR proteins or fragments thereof, in addition to providing the *L. cuprina, M. persicae,* and *B. tabaci* EcR partner protein (USP polypeptide) subunits of ecdysone receptors, and the *L. cuprina, M. persicae,* and *B. tabaci* USP polypeptides of the juvenile hormone receptors of these insects. The present invention further relates to the production of functional recombinant insect receptors and recombinant polypeptide subunits thereof and derivatives and analogues thereof. The present invention further relates to the uses of the recombinant receptor and isolated nucleic acid molecules of the present invention in the regulation of gene expression. The present invention further relates to screening systems and methods of identifying insecticidally-active agents which are capable of agonising or antagonising insect receptor function, such as molecules or ligands which associate with steroid receptors or juvenile hormone receptors so as to modify the affinity of said receptors for their cognate *cis*-acting response elements (eg. insect steroid response elements, juvenile hormone response elements) in the genes which they regulate, or alternatively or in addition, which modify the affinity of said receptors for their cellular stimuli (eg. insect steroids or juvenile hormones) or analogues thereof, or alternatively or in addition, which act as insecticides by virtue of their ability to agonise or antagonise the activity of insect hormones, such as by mimicry of a ligand which binds to said receptor or a ligand-binding region thereof. The invention further extends to such compounds or ligands.

GENERAL

This specification contains nucleotide and amino acid sequence information prepared using the programme PatentIn Version 2.0, presented herein after the bibliography. Each nucleotide or amino acid sequence is identified in the sequence listing by the numeric indicator <210> followed by the sequence identifier (e.g. <210>1, <210>2, etc). The length, type of sequence (DNA, protein (PRT), etc) and source organism for each nucleotide or amino acid sequence are indicated by information provided in the numeric indicator fields <211>, <212> and <213>, respectively. Nucleotide and amino acid sequences referred to in the specification are defined by descriptor "SEQ ID NO:" followed by the numeric identifier. For example, SEQ ID NO: 1 refers to the information provided in the numeric indicator field designated <400> 1, etc.

The designation of nucleotide residues referred to herein are those recommended by the IUPAC-IUB Biochemical Nomenclature Commission, wherein A represents Adenine, C represents Cytosine, G represents Guanine, T represents thymine, Y represents a pyrimidine residue, R represents a purine residue, M represents Adenine or Cytosine, K represents Guanine or Thymine, S represents Guanine or Cytosine, W represents Adenine or Thymine, H represents a nucleotide other than Guanine, B represents a nucleotide other than Adenine, V represents a nucleotide other than Thymine, D represents a nucleotide other than Cytosine and N represents any nucleotide residue.

Bibliographic details of the publications referred to in this specification are collected at the end of the description. Reference herein to prior art, including any one or more prior art documents, is not to be taken as an acknowledgment, or suggestion, that said prior art is common general knowledge in Australia or forms a part of the common general knowledge in Australia.

Throughout this specification, unless the context requires otherwise, the word "comprise", or variations such as "comprises" or "comprising", will be understood to imply the inclusion of a stated step or element or integer or group of steps or elements or integers but not the exclusion of any other step or element or integer or group of elements or integers.

As used herein the term "derived from" shall be taken to indicate that a specified integer may be obtained from a particular source albeit not necessarily directly from that source.

Those skilled in the art will appreciate that the invention described herein is susceptible to variations and modifications other than those specifically described. It is to be understood that the invention includes all such variations and modifications. The invention also includes all of the steps, features, compositions and compounds referred to or indicated in this specification, individually or collectively, and any and all combinations or any two or more of said steps or features.

The present invention is not to be limited in scope by the specific embodiments described herein, which are intended for the purposes of exemplification only. Functionally-equivalent products, compositions and methods are clearly within the scope of the invention, as described herein.

BACKGROUND TO THE INVENTION

International Patent Application No WO91/13167 (applicant, The Board of Trustees of Leyland Stanford University, and hereinafter referred to as WO91/13167) describes the identification, characterization, expression and uses of insect steroid receptors and DNA sequences encoding same and, in particular, the identification, characterization, expression and uses of the steroid receptor of the common fruit fly, *Drosophila melanogaster.*

It has been found by the present inventors that the limited homology between the *D. melanogaster* steroid receptor-encoding gene sequences and the steroid receptor -encoding sequences derived from other insects, in particular those derived from diptera such as the Australian sheep blowfly *L. cuprina*; hemiptera such as the aphid *M. persicae*, leaf sucking insects such as the whitefly *(B. tabaci),* scale insects and leaf hoppers; coleoptera; neuroptera; lepidoptera; and ants, as well as from helminths and protozoa, prevents the routine isolation of DNA sequences encoding steroid receptors or juvenile hormone receptors from these latter-mentioned organisms.

Moreover, the present inventors have discovered that the *D. melanogaster* steroid receptor described in WO91/13167 is temperature-sensitive, showing reduced activity at temperatures above 30°C, such as at temperatures about 37°C, particularly at low concentrations of the

receptor. Accordingly, the *D. melanogaster* steroid receptor described in WO91/13167 is of limited utility at physiological temperatures applicable to animal or bacterial cells. Moreover, wherein it is desirable to produce a biologically-active steroid receptor using *in vivo* or *in situ* expression systems, which expression systems routinely utilise cells or tissues in the temperature range of about 28°C to about 42°C, the *D. melanogaster* steroid receptor is also of limited utility.

In work leading up to the present invention, the present inventors developed a novel screening protocol, involving the utilisation of highly-degenerate oligonucleotide probes and primers derived from the amino acid sequences of the DNA-binding domains of the *D. melanogaster* and *Chironomus tentans* ecdysone receptor polypeptides, to identify nucleotide sequences encoding novel steroid receptor polypeptides and novel insect juvenile hormone receptor polypeptides. The present inventors have further identified specific regions within these novel polypeptides which are suitable for use in preparing a surprising range of novel steroid receptor polypeptide derivatives and insect juvenile hormone receptor polypeptide derivatives. The novel steroid receptor polypeptides and novel insect juvenile hormone receptor polypeptides of the present invention, and derivative polypeptides thereof, and assembled steroid receptors and insect juvenile hormone receptors derived from said polypeptides and derivatives, and nucleic acid molecules encoding same as exemplified herein, provide the means for developing a wide range of insecticidally-active agents, as well as methods for the regulated production of bioactive molecules. In particular, the present invention provides the means for developing specific ligands which bind to and either agonise or antagonise the steroid receptors or juvenile hormone receptors, and/or which bind to polypeptide subunits of said receptors as described herein, thereby functioning as highly-specific insecticides, offering significant commercial and environmental benefits.

The present inventors have been surprisingly successful in characterizing the ecdysone receptor and juvenile hormone receptor derived from insects of the orders Diptera and Hemiptera, and polypeptide components thereof and functional derivatives of said polypeptides and receptors, particularly in light of the extreme difficulties in dealing with these organisms. The nature of these molecules was unknown prior to the present invention.

The various aspects of this invention overcome the problems associated with *Drosophila* ecdysone receptors which lack thermal stability. Moreover, those aspects of the invention pertaining to methods of screening for insecticidally active agents do not involve competition assays which are generally complex, and often inaccurate or difficult to calibrate.

SUMMARY OF THE INVENTION

One aspect of the present invention provides an isolated nucleic acid molecule comprising a nucleotide sequence which encodes or is complementary to a sequence which encodes a steroid receptor polypeptide or a juvenile hormone receptor polypeptide or a bioactive derivative or analogue thereof, wherein said polypeptide:

(i) is selected from the group consisting of the EcR polypeptide of an steroid receptor, the partner protein (USP polypeptide) of an steroid receptor and the USP polypeptide of a juvenile hormone receptor; and

(ii) comprises an amino acid sequence that is at least 40% identical to an amino acid sequence selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 38, SEQ ID NO: 40, and SEQ ID NO: 42.

In an alternative embodiment, the present invention provides an isolated nucleic acid molecule comprising a nucleotide sequence which encodes or is complementary to a sequence which encodes a steroid receptor polypeptide or a juvenile hormone receptor polypeptide or a bioactive derivative or analogue thereof, wherein said polypeptide:

(i) is selected from the group consisting of the EcR polypeptide of an steroid receptor, the partner protein (USP polypeptide) of an steroid receptor and the USP polypeptide of a juvenile hormone receptor; and

(ii) comprises an amino acid sequence that is at least 40% identical to an amino acid sequence encoded by the DNA of insects which is present in a plasmid selected from the group consisting of AGAL Accession Nos. NM99/04565, NM99/04566, NM99/04567, NM99/04568, NM00/12580, and NM00/12581.

In a further alternative embodiment, the isolated nucleic acid molecule of the present invention comprises a nucleotide sequence which encodes or is complementary to a sequence which encodes a steroid receptor polypeptide or a juvenile hormone receptor polypeptide or a bioactive derivative or analogue thereof, wherein said nucleotide sequence is selected from the group consisting of:

(i) a nucleotide sequence having at least 40% identity to a nucleotide sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 37, SEQ ID NO: 39, and SEQ ID NO: 41, or a complementary nucleotide sequence to any one of said sequences;

(ii) a nucleotide sequence that is capable of hybridising under at least low stringency conditions to a nucleotide sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 37, SEQ ID NO: 39, and SEQ ID NO: 41, or to a complementary nucleotide sequence to any one of said sequences;

(iii) a nucleotide sequence that is capable of hybridising under at least low stringency conditions to a nucleotide sequence contained in a plasmid selected from the group consisting of AGAL Accession Nos. NM99/04565, NM99/04566, NM99/04567, NM99/04568, NM00/12580, and NM00/12581;

(iv) a nucleotide sequence having at least 40% identity to a nucleotide sequence contained in a plasmid selected from the group consisting of AGAL Accession Nos. NM99/04565, NM99/04566, NM99/04567, NM99/04568, NM00/12580, and NM00/12581; and

(v) a nucleotide sequence that is amplifiable by PCR using a nucleic acid primer sequence selected from the group consisting of SEQ ID NO: 23, SEQ ID NO: 24, SEQ ID NO: 25, SEQ ID NO: 26, SEQ ID NO: 27, SEQ ID NO: 28, SEQ ID NO: 31, and SEQ ID NO: 32.

In a further alternative embodiment, the present invention provides an isolated nucleic acid molecule which encodes a steroid receptor EcR polypeptide and comprises the nucleotide

sequence set forth in SEQ ID NO: 1, or SEQ ID NO: 13.

In a further alternative embodiment, the present invention provides an isolated nucleic acid molecule which encodes a steroid receptor USP polypeptide or a juvenile hormone receptor polypeptide and comprises a nucleotide sequence selected from the group consisting of SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 37, SEQ ID NO: 39, and SEQ ID NO: 41.

In a further alternative embodiment, the present invention provides an isolated nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 21, SEQ ID NO: 37, and SEQ ID NO: 41, or a complementary nucleotide sequence to any one of said sequences.

In a further alternative embodiment, the present invention provides an isolated nucleic acid molecule which comprises the nucleotide sequence selected from the group consisting of SEQ ID NO: 23, SEQ ID NO: 24, SEQ ID NO: 25, SEQ ID NO: 26, SEQ ID NO: 27, SEQ ID NO: 28, SEQ ID NO: 31, and SEQ ID NO: 32, or a complementary nucleotide sequence thereto.

A second aspect of the present invention provides a method of identifying an isolated nucleic acid molecule which encodes a steroid receptor polypeptide or a juvenile hormone receptor polypeptide comprising the steps of:

(i) hybridising genomic DNA, mRNA or cDNA with a hybridisation-effective amount of one or more probes selected from the group consisting of:

(a) a probe comprising at least 10 contiguous nucleotides in length from a nucleotide sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 24, SEQ ID NO: 25, SEQ ID NO: 26, SEQ ID NO: 27, SEQ ID NO: 28, SEQ ID NO: 31, SEQ ID NO: 32, SEQ ID NO: 37, SEQ ID NO: 39, and SEQ ID NO: 41, or a complementary nucleotide sequence to any one of said sequences;

(b) a probe comprising at least 10 contiguous nucleotides in length from cDNA contained in a plasmid selected from the group consisting of AGAL Accession Nos. NM99/04565, NM99/04566, NM99/04567, NM99/04568, NM00/12580, and NM00/12581; and

(c) a hybridisation probe comprising a nucleotide sequences selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 37, SEQ ID NO: 39, and SEQ ID NO: 41, or a complementary nucleotide sequence to any one of said sequences, or a homologue, analogue or derivative of any one of said sequences or complementary sequences having at least 40% identity thereto; and

(ii) detecting the hybridisation.

In an alternative embodiment, the inventive method comprises the steps of:

(i) annealing to genomic DNA, mRNA or cDNA, one or more PCR primers selected from the group consisting of:

(a) a primer comprising at least 10 contiguous nucleotides in length from a nucleotide sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 24, SEQ ID NO: 25, SEQ ID NO: 26, SEQ ID NO: 27, SEQ ID NO: 28, SEQ ID NO: 31, SEQ ID NO: 32, SEQ ID NO: 37, SEQ ID NO: 39, and SEQ ID NO: 41, or a complementary nucleotide sequence to any one of said sequences;

(b) a primer comprising at least 10 contiguous nucleotides in length from cDNA contained in a plasmid selected from the group consisting of AGAL Accession Nos. NM99/04565, NM99/04566, NM99/04567, NM99/04568, NM00/12580, and NM00/12581; and

(ii) amplifying a nucleotide sequence which encodes a steroid receptor polypeptide or a juvenile hormone receptor polypeptide in a polymerase chain reaction.

In a further alternative embodiment, the inventive method comprises the steps of:

(i) amplifying a nucleotide sequence which encodes a steroid receptor polypeptide or a juvenile hormone receptor polypeptide in a polymerase chain reaction using one or more PCR primers selected from the group consisting of:

(a) a primer comprising at least 10 contiguous nucleotides in length from a nucleotide sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 24, SEQ ID NO: 25, SEQ ID NO: 26, SEQ ID NO: 27, SEQ ID NO: 28, SEQ ID NO: 31, SEQ ID NO: 32, SEQ ID NO: 37, SEQ ID NO: 39, and SEQ ID NO: 41, or a complementary nucleotide sequence to any one of said sequences; and

(b) a primer comprising at least 10 contiguous nucleotides in length from cDNA contained in a plasmid selected from the group consisting of AGAL Accession Nos. NM99/04565, NM99/04566, NM99/04567, NM99/04568, NM00/12580, and NM00/12581;

(ii) hybridising the amplified nucleotide sequence to genomic DNA, mRNA or cDNA with a hybridisation-effective amount of one or more probes selected from the group consisting of:

(a) a probe comprising at least 10 contiguous nucleotides in length derived from a sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 24, SEQ ID NO: 25, SEQ ID NO: 26, SEQ ID NO: 27, SEQ ID NO: 28, SEQ ID NO: 31, SEQ ID NO: 32, SEQ ID NO: 37, SEQ ID NO: 39, and SEQ ID NO: 41, or a complementary nucleotide sequence to any one of said sequences;

(b) a probe comprising at least 10 contiguous nucleotides in length derived from a cDNA contained in a plasmid selected from the group consisting of AGAL Accession Nos. NM99/04565, NM99/04566, NM99/04567, NM99/04568,

NM00/12580, and NM00/12581; and

(c) a hybridisation probe comprising a nucleotide sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 37, SEQ ID NO: 39, and SEQ ID NO: 41, or a complementary nucleotide sequence to any one of said sequences, or a homologue, analogue or derivative of any one of said sequences or complementary sequences having at least 40% identity thereto; and

(iii) detecting the hybridisation.

A third aspect of the present invention provides a genetic construct comprising the subject isolated nucleic acid molecule which encodes a steroid receptor polypeptide or a juvenile hormone receptor polypeptide, operably linked to a promoter sequence. Preferably, the subject nucleic acid molecule is in an expressible format, such that it is possible to produce a recombinant polypeptide therefrom.

Accordingly, a fourth aspect of the invention provides a recombinant or isolated polypeptide comprising a steroid receptor polypeptide or juvenile hormone receptor polypeptide or a bioactive derivative or analogue thereof, wherein said polypeptide:

(i) is selected from the group consisting of the EcR polypeptide of a steroid receptor, the partner protein (USP polypeptide) of a steroid receptor and the USP polypeptide of a juvenile hormone receptor; and

(ii) comprises an amino acid sequence that is at least 40% identical to an amino acid sequence selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 38, SEQ ID NO: 40, and SEQ ID NO: 42;

wherein said polypeptide is substantially free of naturally-associated insect cell components.

In an alternative embodiment, the invention provides a recombinant or isolated polypeptide

comprising a steroid receptor polypeptide or juvenile hormone receptor polypeptide or a bioactive derivative or analogue thereof, wherein said polypeptide:

(i) is selected from the group consisting of the EcR polypeptide of a steroid receptor, the partner protein (USP polypeptide) of a steroid receptor and the USP polypeptide of a juvenile hormone receptor; and

(ii) comprises an amino acid sequence that is at least 40% identical to a polypeptide encoded by cDNA present in a plasmid selected from the group consisting of AGAL Accession Nos. NM99/04565, NM99/04566, NM99/04567, NM99/04568, NM00/12580, and NM00/12581;

wherein said polypeptide is substantially free of naturally-associated insect cell components.

A fifth aspect of the invention provides a cell comprising the subject isolated nucleic acid molecule which encodes a steroid receptor polypeptide or a juvenile hormone receptor polypeptide.

In a preferred embodiment, the cell of the present invention expresses the polypeptide encoded by the nucleic acid molecule.

In a preferred embodiment, the cell expresses a steroid receptor polypeptide or a fragment thereof which receptor is capable of binding to an insect steroid or analogue thereof or a candidate insecticidally active agent to form an activated complex, and comprises a nucleic acid sequence encoding a bioactive molecule or a reporter molecule operably linked to one or more insect steroid response elements which on binding of the said activated complex promotes transcription of the nucleic acid sequence, wherein said cell on exposure to insect steroid or an analogue thereof, regulates expression of said bioactive molecule or allows detection of said reporter molecule.

In a further aspect of this invention, there is provided an animal (such as a mammal), microorganism, plant or aquatic organism, containing one or more cells as mentioned above.

A further aspect of the present invention provides a method of identifying a modulator of insect

steroid receptor-mediated gene expression or insect juvenile hormone receptor-mediated gene expression comprising:

(i) assaying the expression of a reporter gene in the presence of a recombinant or isolated insect steroid receptor polypeptide or a juvenile hormone receptor polypeptide of the invention and a potential modulator; and

(ii) assaying the expression of a reporter gene in the presence of a recombinant or isolated insect steroid receptor polypeptide or a juvenile hormone receptor polypeptide of the invention and without said potential modulator; and

(ii) comparing expression of the reporter gene in the presence of the potential modulator to the expression of a reporter gene in the absence of the potential modulator,

wherein said reporter gene is placed operably under the control of a steroid response element (SRE) to which said insect steroid receptor binds or a promoter sequence comprising said SRE.

A still further aspect of the invention provides a method of identifying a potential insecticidal compound comprising:

(i) assaying the binding directly or indirectly of a recombinant or isolated insect steroid receptor polypeptide or a juvenile hormone receptor polypeptide of the invention to a steroid response element (SRE) to which said insect steroid receptor binds, in the presence of a candidate compound; and

(ii) assaying the binding directly or indirectly of a recombinant or isolated insect steroid receptor polypeptide or a juvenile hormone receptor polypeptide of the invention to a steroid response element (SRE) to which said insect steroid receptor binds, in the absence of said candidate compound; and

(ii) comparing the binding assayed at (i) and (ii), wherein a difference in the level of binding indicates that the candidate compound possesses potential insecticidal activity.

A still further aspect of the invention provides a method of identifying a candidate insecticidally-active agent comprising the steps of:

a) expressing an EcR polypeptide of a steroid receptor or a fragment thereof which

includes the ligand-binding region, optionally in association with an EcR partner protein (USP polypeptide) of a steroid receptor or ligand binding domain thereof, optionally in association with an insect steroid or analogue thereof so as to form a complex;

b) purifying or precipitating the complex;

c) determining the three-dimensional structure of the ligand binding domain of the complex; and

d) identifying compounds which bind to or associate with the three-dimensional structure of the ligand binding domain, wherein said compounds represent candidate insecticidally-active agents.

A still further aspect of the invention provides a method of identifying a candidate insecticidally-active agent comprising the steps of:

a) expressing a USP polypeptide of a juvenile hormone receptor or a fragment thereof which includes the ligand-binding region, optionally in association with an EcR polypeptide of a steroid receptor or ligand binding domain thereof, and optionally in association with an insect steroid or analogue thereof, so as to form a complex;

b) purifying or precipitating the complex;

c) determining the three-dimensional structure of the ligand binding domain of the complex; and

d) identifying compounds which bind to or associate with the three-dimensional structure of the ligand binding domain, wherein said compounds represent candidate insecticidally-active agents.

In another aspect this invention relates to a method or assay for screening insecticidally active compounds which comprises reacting a candidate insecticidal compound with a steroid receptor polypeptide or fragment thereof encompassing the ligand binding domain, or complex thereof with a partner protein or a fragment thereof which encompasses the ligand binding domain, and detecting binding or absence of binding of said compound so as to determine insecticidal activity.

A still further aspect of the invention provides a synthetic compound which interacts with the

three dimensional structure of a polypeptide or protein selected from the group consisting of:

(i) an EcR polypeptide of a steroid receptor or a fragment thereof;

(ii) an EcR partner protein (USP polypeptide) of a steroid receptor or a fragment thereof;

(iii) a USP polypeptide of a juvenile hormone receptor; and

(iv) a functional receptor or protein complex formed by association of (i) and (ii),

wherein said compound is capable of binding to said polypeptide or protein to agonise or antagonise the binding activity or bioactivity thereof.

Preferably, the synthetic compounds are derived from the three dimensional structure of insect steroid receptor(s) or juvenile hormone receptor(s) which compounds bind to said receptor(s) and have the effect of either inactivating the receptor(s) or potentiating the activity of the receptor(s). More preferably, the compounds mimic the three-dimensional structure of a ligand which binds to the receptor(s) and more preferably, mimic the three-dimensional structure of a ligand which binds to the ligand-binding region of said receptor(s).

In a still further aspect of this invention, there is provided a screening system for insecticidally active agents comprising a nucleotide sequence encoding a steroid receptor or a fragment thereof, and a nucleotide sequence encoding a partner protein or a fragment thereof which associates with the receptor so as to confer enhanced affinity for insect steroid response elements, or enhanced affinity for insect steroids or analogues thereof or insecticidally active agents, or thermostability or enhanced thermostability of said receptor, which receptor and partner protein is capable of binding to a candidate insecticidally active agent to form an activated complex, and a nucleic acid sequence encoding a bioactive molecule or a reporter molecule operably linked to one or more insect steroid response elements which on binding of the said activated complex regulates transcription of the nucleic acid sequence, wherein on exposure to said agent expression of the bioactive molecule or reporter molecule correlates with insecticidal activity.

In another aspect of this invention, there is provided a method for the regulated production of a bioactive molecule or a reporter molecule in a cell, said method comprising the steps of

introducing into said cell:

a) a nucleotide sequence encoding a steroid receptor or a fragment thereof which is capable of binding an insect steroid or analogue thereof, to form an activated complex; and

b) a nucleotide sequence encoding said bioactive molecule or reporter molecule operably linked to one or more insect steroid response elements which on binding of the said activated complex regulates transcription of the nucleic acid sequence encoding said bioactive molecule or reporter molecule,

wherein exposing the cell to an insect steroid or analogue thereof regulates expression of the bioactive molecule or reporter molecule.

SUMMARY OF SEQUENCE LISTING

SEQ ID NO: 1:	The nucleotide sequence of the open reading frame of a cDNA molecule which encodes the EcR polypeptide subunit of the *L. cuprina* ecdysone receptor and amino acid sequence therefor.
SEQ ID NO: 2:	The amino acid sequence of the EcR polypeptide subunit of the *L. cuprina* ecdysone receptor.
SEQ ID NO: 3:	The nucleotide sequence of the cDNA molecule contained in plasmid pBLU1 which encodes the EcR partner protein (USP polypeptide) subunit of the *L. cuprina* ecdysone receptor or which encodes the USP polypeptide subunit of the *L. cuprina* juvenile hormone receptor, and amino acid sequence therefor.
SEQ ID NO: 4:	The amino acid sequence of the EcR partner protein (USP polypeptide) subunit of the *L. cuprina* ecdysone receptor or the amino acid sequence of the USP polypeptide subunit of the *L. cuprina* juvenile hormone receptor, encoded by SEQ ID NO: 3.
SEQ ID NO: 5:	The nucleotide sequence of the cDNA molecule from plasmid pLSP5 which encodes the EcR partner protein (USP polypeptide) subunit of the *L. cuprina* ecdysone receptor or which encodes the USP polypeptide subunit of the *L. cuprina* juvenile hormone receptor, and amino acid sequence therefor.

SEQ ID NO: 6: The amino acid sequence of the EcR partner protein (USP polypeptide) subunit of the *L. cuprina* ecdysone receptor or the amino acid sequence of the USP polypeptide subunit of the *L. cuprina* juvenile hormone receptor, encoded by SEQ ID NO: 5.

SEQ ID NO: 7: The nucleotide sequence of the cDNA molecule from plasmid pLSP12 which encodes the EcR partner protein (USP polypeptide) subunit of the *L. cuprina* ecdysone receptor or which encodes the USP polypeptide subunit of the *L. cuprina* juvenile hormone receptor, and amino acid sequence therefor.

SEQ ID NO: 8: The amino acid sequence of the EcR partner protein (USP polypeptide) subunit of the *L. cuprina* ecdysone receptor or the amino acid sequence of the USP polypeptide subunit of the *L. cuprina* juvenile hormone receptor, encoded by SEQ ID NO: 7.

SEQ ID NO: 9: The nucleotide sequence of a cDNA molecule which encodes part of the EcR polypeptide subunit of the *M. persicae* ecdysone receptor and amino acid sequence therefor.

SEQ ID NO: 10: The amino acid sequence of a part of the EcR polypeptide subunit of the *M. persicae* ecdysone receptor.

SEQ ID NO: 11: The nucleotide sequence of the EcR probe 1 which is specific for genetic sequences encoding the EcR polypeptide subunit of aphid ecdysone receptors, in particular the EcR polypeptide subunit of the *M. persicae* ecdysone receptor.

SEQ ID NO: 12: The nucleotide sequence of the EcR probe 2 sequence which is specific for genetic sequences encoding the EcR polypeptide subunit of aphid ecdysone receptors, in particular the EcR polypeptide subunit of the *M. persicae* ecdysone receptor.

SEQ ID NO: 13: The nucleotide sequence of the open reading frame of a cDNA molecule which encodes the EcR polypeptide subunit of the *M. persicae* ecdysone receptor and amino acid sequence therefor.

SEQ ID NO: 14: The amino acid sequence of the EcR polypeptide subunit of the *M. persicae* ecdysone receptor.

SEQ ID NO: 15: The nucleotide sequence of the open reading frame of a first cDNA molecule encoding the EcR partner protein (USP polypeptide) subunit of the *M. persicae* ecdysone receptor or the USP polypeptide subunit of the *M. persicae* juvenile hormone receptor, and amino acid sequence therefor.

SEQ ID NO: 16: The amino acid sequence of the EcR partner protein (USP polypeptide) subunit of the *M. persicae* ecdysone receptor or the amino acid sequence of the USP polypeptide subunit of the *M. persicae* juvenile hormone receptor, encoded by SEQ ID NO: 15.

SEQ ID NO: 17: The nucleotide sequence of the open reading frame of a second cDNA molecule encoding the EcR partner protein (USP polypeptide) subunit of the *M. persicae* ecdysone receptor or the USP polypeptide subunit of the *M. persicae* juvenile hormone receptor, and amino acid sequence therefor.

SEQ ID NO: 18: The amino acid sequence of the EcR partner protein (USP polypeptide) subunit of the *M. persicae* ecdysone receptor or the amino acid sequence of the USP polypeptide subunit of the *M. persicae* juvenile hormone receptor, encoded by SEQ ID NO: 17.

SEQ ID NO: 19: The nucleotide sequence of the open reading frame of a third cDNA molecule encoding the EcR partner protein (USP polypeptide) subunit of the *M. persicae* ecdysone receptor or the USP polypeptide subunit of the *M. persicae* juvenile hormone receptor, and amino acid sequence therefor.

SEQ ID NO: 20: The amino acid sequence of the EcR partner protein (USP polypeptide) subunit of the *M. persicae* ecdysone receptor or the amino acid sequence of the USP polypeptide subunit of the *M. persicae* juvenile hormone receptor, encoded by SEQ ID NO: 19.

SEQ ID NO: 21: The nucleotide sequence of a 150 base-pair probe which is specific for genetic sequences encoding the EcR partner protein (USP polypeptide) subunit of *L. cuprina* ecdysone receptor or the USP polypeptide subunit of the *L. cuprina* juvenile hormone receptor, and amino acid sequence

therefor.

SEQ ID NO: 22: The amino acid sequence encoded by the nucleotide sequence of SEQ ID NO: 21, comprising amino acid residues 108-149 of the EcR partner protein (USP polypeptide) subunit of the *L. cuprina* ecdysone receptor or amino acid residues 108-149 of the amino acid sequence of the USP polypeptide subunit of the *L. cuprina* juvenile hormone receptor set forth herein as SEQ ID NO: 4.

SEQ ID NO: 23: The nucleotide sequence of the degenerate primer Rdna3.

SEQ ID NO: 24: The nucleotide sequence of the degenerate primer Rdna4.

SEQ ID NO: 25: The nucleotide sequence of the primer Mdna1.

SEQ ID NO: 26: The nucleotide sequence of the primer Mdna2.

SEQ ID NO: 27: The nucleotide sequence of the primer AP1.

SEQ ID NO: 28: The nucleotide sequence of the degenerate primer AP2.

SEQ ID NO: 29: The sequence of oligonucleotide SPX5 used to construct plasmid pVPLcEcR.

SEQ ID NO: 30: The sequence of oligonucleotide XPS5 used to construct plasmid pVPLcEcR.

SEQ ID NO: 31: The nucleotide sequence of oligonucleotide A used to construct plasmid pSGDM.

SEQ ID NO: 32: The nucleotide sequence of oligonucleotide B used to construct plasmid pSGDM.

SEQ ID NO: 33: The sequence of oligonucleotide A used to construct the expression plasmid pMpEcR.LcUSP.DUAL.

SEQ ID NO: 34: The sequence of oligonucleotide B used to construct the expression plasmid pMpEcR.LcUSP.DUAL.

SEQ ID NO: 35: The sequence of oligonucleotide C used to construct the expression plasmid pMpEcR.USP.DUAL.

SEQ ID NO: 36: The sequence of oligonucleotide D used to construct the expression plasmid pMpEcR.USP.DUAL.

SEQ ID NO: 37: The nucleotide sequence of a probe which is specific for genetic sequences encoding the EcR partner protein (USP polypeptide) subunit

of *B. tabaci* ecdysone receptor or the USP polypeptide subunit of the *B. tabaci* juvenile hormone receptor, and amino acid sequence therefor.

SEQ ID NO: 38: The amino acid sequence encoded by the nucleotide sequence of SEQ ID NO: 37.

SEQ ID NO: 39: The nucleotide sequence of the open reading frame of a cDNA molecule encoding the EcR partner protein (USP polypeptide) subunit of the *B. tabaci* ecdysone receptor or the USP polypeptide subunit of the *B. tabaci* juvenile hormone receptor, and amino acid sequence therefor.

SEQ ID NO: 40: The amino acid sequence of the EcR partner protein (USP polypeptide) subunit of the *B. tabaci* ecdysone receptor or the amino acid sequence of the USP polypeptide subunit of the *B. tabaci* juvenile hormone receptor encoded by SEQ ID NO: 39.

SEQ ID NO: 41: The nucleotide sequence of a probe which is specific for genetic sequences encoding the EcR polypeptide subunit of *B. tabaci* ecdysone receptor, and amino acid sequence therefor.

SEQ ID NO: 42: The amino acid sequence encoded by the nucleotide sequence of SEQ ID NO: 37.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is a graphical representation showing function of the EcR polypeptide subunit of the *L. cuprina* ecdysone receptor *in vivo*. CHO cells were co transfected with:

(1) one of the following expression plasmids: pSGDmEcR, pSGLcEcR, or the parental expression plasmid pSG5 as a control, at 1μg/ml;

(2) plasmid p(EcRE)$_7$-CAT (1 μg/ml); and

(3) an independent reporter plasmid , pPGKLacZ, at 1 μg/ml.

CAT expression was induced with Muristerone A at either 10 μM or 50 μM while control cells received only the carrier ethanol. ELISA kits were used to quantify the synthesis of CAT and β-galactosidase in extracts of cells forty eight hours after transfection. The level of CAT was normalized to the level of β-galactosidase in the same extract. Fold-induction represents the normalized values for CAT gene expression in cells transfected with pSGDmEcR, pSGLcEcR or pSG5 in the presence of hormone, relative to the normalized values for CAT gene

expression in cells transfected with the same plasmid, but in the absence of hormone. The average values of three independent experiments are shown and the error bars indicate standard error of the mean.

Figure 2 is a copy of a graphical representation showing the activity of plasmids pSGLD and pSGDL, containing chimeric EcR polypeptide subunits of insect ecdysone receptors, produced as described in the Examples. Cotransfection assays were performed as described in the Examples using plasmids pSGLD and pSGDL and the CAT reporter plasmid p(EcRE)$_7$ -CAT (1ug/ml) and an independent reporter, pPGKLacZ at 1 µg/ml each. CAT/β-Gal (%) refers to CAT reporter activity expressed as a percentage relative to β-galactosidase activity produced by the internal control reporter, pPGKLacZ.

Figure 3 is a copy of a graphical representation showing the binding activity in extracts of Sf9 and Sf21 cells containing a baculovirus expressing LcEcRDEF and LcUSPDEF, as described in the Examples. Control cells contained baculovirus expressing β-glucuronidase and CAT only.

Figure 4 is a graphical representation showing the ecdysteroid binding activities of an *in vitro*-translated *Myzus persicae* EcR (MpEcR) polypeptide, an *in vitro*-translated *Myzus persicae* USP (MpUSP) polypeptide, and a complex formed by *in vitro*-translated *M. persicae* EcR and USP polypeptides.

Figure 5 is a copy of a graphical representation showing the expression activity of plasmid pVPLcEcR, encoding a chimeric *L. cuprina* EcR polypeptide, and plasmid pSGLcUSP encoding the *L. cuprina* EcR partner protein (USP polypeptide), in CV1 cells, in accordance with the description provided in Example 18. The CAT reporter plasmid p(EcRE)$_7$ -CAT (1ug/ml), and an independent reporter plasmid, pPGKLacZ (1 µg/ml) were used to assay ecdysteroid-dependent gene expression. Data indicate expression of the CAT reporter gene relative to the level of expression of the transfection control β-galactosidase reporter gene. The symbols + and - indicate the presence or absence, respectively, of the plasmids pVPLcEcR and pSGLcUSP, or the presence (+) or absence (-) of 1 µM Ponasterone A (PonA). Error bars

indicate the standard error of the mean.

Figure 6 is a graphical representation showing *in vivo* function of a modified EcR polypeptide subunit of the *M. persicae* ecdysone receptor in CHO cells. The CHO cells were co-transfected with a reporter plasmid p(EcRE)$_7$-CAT (1µg/ml) and an expression plasmid selected from the group consisting of pSGDmEcR, pSGMpEcR, pSGDM, pSGMD, and pSG5 (described in the examples), also at 1µg/ml concentration. Data indicate CAT reporter gene expression as determined by ELISA, for cells lacking Muristerone A (open bars) or containing 10µM Muristerone A (filled bars). The level of CAT expression is directly correlated to the concentration of the product of the enzymatic reaction in the assay and was measured as an absorbance at 405nm.

Figure 7 is a copy of a graphical representation showing the binding of [^{3}H] ponasterone A to extracts of Sf9 cells infected with baculovirus expressing the ligand binding regions (i.e. domains D/E/F) of (i) the *M. persicae* EcR polypeptide and the *L. cuprina* EcR partner protein (USP polypeptide); (ii) the *M. persicae* EcR polypeptide and the *M. persicae* EcR partner protein (USP polypeptide); and (iii) the *L. cuprina* EcR polypeptide and the *L. cuprina* EcR partner protein (USP polypeptide). Highly significant binding (i.e. above background) of the ecdysteroid analogue is apparent for all three constructs tested.

Figure 8 is a copy of a graphical representation showing the activity of plasmid pSGDM (Example 19), encoding a chimeric *M. persicae* EcR polypeptide, and plasmid pBKMpUSP1, encoding an *M. persicae* EcR partner protein (USP polypeptide), in CV1 cells. The CAT reporter plasmid p(EcRE)$_7$ -CAT (1ug/ml) and an internal control reporter plasmid, pPopNLacZ (1 µg/ml) were present in all assays. The symbols + and - indicate the presence or absence, respectively, of plasmids indicated in the figure, or the presence (+) or absence (-) of 10 µM Ponasterone a. Data indicate expression of the CAT reporter gene relative to the level of expression of the independent reporter gene β-galactosidase. Error bars indicate the standard error of the mean.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

One aspect of the present invention provides an isolated nucleic acid molecule comprising a nucleotide sequence which encodes or is complementary to a sequence which encodes a steroid receptor polypeptide or a juvenile hormone receptor polypeptide or a bioactive derivative or analogue thereof, wherein said polypeptide:

(i) is selected from the group consisting of the EcR polypeptide of a steroid receptor, the partner protein (USP polypeptide) of a steroid receptor and the USP polypeptide of a juvenile hormone receptor; and

(ii) comprises an amino acid sequence having at least 40% identity to an amino acid sequence selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 38, SEQ ID NO: 40, and SEQ ID NO: 42.

Accordingly, the isolated nucleic acid molecule of the invention may comprise a fragment of a nucleotide sequence encoding a full-length receptor polypeptide.

It is to be understood that a "fragment" of a nucleotide sequence encoding an EcR polypeptide subunit of a steroid receptor or an EcR partner protein (USP polypeptide) of a steroid receptor or a USP polypeptide of a juvenile hormone receptor, refers to a nucleotide sequence encoding a part or fragment of such a receptor which is capable of binding or associating with an insect steroid or an analogue thereof, or a candidate insecticidally active compound. Fragments of a nucleotide sequence would generally comprise in excess of twenty contiguous nucleotides derived from the base sequence and may encode one or more domains of a functional insect steroid receptor or juvenile hormone receptor.

Preferably, the isolated nucleic acid molecule of the invention encodes an ecdysteroid receptor polypeptide. Those skilled in the art are aware that ecdysteroid receptors derived from insects are heterodimeric receptors comprising an EcR polypeptide subunit and an EcR partner protein (USP polypeptide) (*see* also Jones and Sharp, 1997). In this regard, the present inventors have discovered that the USP polypeptide of the insect juvenile hormone receptor is structurally-

identical to the EcR partner protein of the ecdysteroid receptor of the present invention, however juvenile hormone receptors comprise monomers or multimers of the USP polypeptide acting without the EcR polypeptide subunit that is present in the ecdysteroid receptors. Accordingly, the present invention extends equally to nucleotide sequences encoding polypeptides of both the ecdysteroid receptors and polypeptides of the juvenile hormone receptors of insects.

More preferably, the isolated nucleic acid molecule of the invention encodes an ecdysteroid receptor that is modulated by one or more of the steroids ecdysone, ponasterone A, or muristerone, or an analogue of an ecdysteroid.

The isolated nucleic acid molecule of the invention may be derived from any organism that contains steroid receptors that are responsive to ecdysteroids or ecdysteroid-like compounds or juvenile hormones, or analogues of such receptor-ligands. Accordingly, the present invention is not to be limited in any of its embodiments to the particular source of the subject nucleic acid, or polypeptide encoded therefor.

Preferably, the isolated nucleic acid molecule of the invention is derived from insects, helminths (nematodes, cestodes, trematodes), protozoa, and ants, amongst others.

More preferably, the isolated nucleic acid molecule of the invention is derived from an insect selected from the group consisting of diptera, hemiptera, coleoptera, neuroptera, lepidoptera and ants, amongst others. Still more preferably, the isolated nucleic acid molecule of the present invention is derived from aphids, scale insects, leaf hoppers, white fly, and blowflies such as sheep blowflies.

The present invention does not extend to amino acid sequences comprising the complete EcR polypeptide subunit of the *D. melanogaster* ecdysone receptor as described in WO91/13167. However, this exclusion is made on the understanding that the present invention does encompass chimeric genes and fusion proteins which include the *D. melanogaster* nucleotide and amino acid sequences, respectively.

In a particularly preferred embodiment, the isolated nucleic acid molecule of the present invention is derived from the aphid *M. persicae* or alternatively, from the Australian sheep blowfly, *L. cuprina*.

The ecdysteroid receptor is preferably modulated by one or more of the steroids ecdysone, ponasterone A, or muristerone, or an analogue of an ecdysteroid.

As used herein, the term "analogue of an ecdysteroid" shall be taken to indicate any compound that binds to one or more polypeptide subunits of an ecdysteroid receptor or the heterodimeric holoreceptor comprising same or alternatively or in addition, which binds to the USP polypeptide of a juvenile hormone receptor or alternatively or in addition, which binds to a bioactive derivative or analogue of said polypeptides or holoreceptor. The term "analogue of an ecdysteroid" shall further be taken to indicate any compound that modulates the bioactivity of one or more polypeptide subunits of an ecdysteroid receptor or the heterodimeric holoreceptor comprising same or alternatively or in addition, that modulates the bioactivity of the USP polypeptide of a juvenile hormone receptor or alternatively or in addition, that modulates the bioactivity of a bioactive derivative or analogue of said polypeptides or holoreceptor.

The present invention is not to be limited in scope to the specific *L. cuprina, M. persicae,* and *B. tabaci* nucleotide and amino acid sequences set forth in the accompanying Sequence Listing, and persons skilled in the art will readily be able to identify additional related sequences from other sources using art-recognised procedures, for example using nucleic acid hybridisation and/or polymerase chain reaction essentially as described by Ausubel *et al.* (1992) and/or McPherson *et al.* (1991) and/or Sambrook *et al* . (1989).

Accordingly, the present invention clearly encompasses isolated nucleic acid molecules which encode or are complementary to isolated nucleic acid molecules which encode the subject EcR polypeptide of a steroid receptor or fragments thereof, or the subject EcR partner protein (USP polypeptide) of a steroid receptor or the subject USP polypeptide of a juvenile hormone receptor, in addition to derivatives, fragments and analogues thereof which comprise amino

acid sequences having at least 40% identity to an amino acid sequence selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 38, SEQ ID NO: 40, and SEQ ID NO: 42.

The present invention clearly extends further to isolated nucleic acid molecules which encode or are complementary to isolated nucleic acid molecules which encode the subject EcR polypeptide of a steroid receptor or fragments thereof, or the subject EcR partner protein (USP polypeptide) of a steroid receptor or the subject USP polypeptide of a juvenile hormone receptor, in addition to derivatives, fragments and analogues thereof which comprise amino acid sequences having at least 40% identity to an amino acid sequence encoded by *L. cuprina, M. persicae* or *B. tabaci* cDNA contained in a plasmid selected from the group consisting of AGAL Accession Nos. NM99/04565, NM99/04566, NM99/04567, NM99/04568, NM00/12580, and NM00/12581.

For the purposes of nomenclature, plasmid pLcEcR contains the cDNA encoding the EcR polypeptide subunit of the *Lucilia cuprina* ecdysone receptor. This plasmid was deposited on 1 July, 1999 with the Australian Government Analytical Laboratories at 1 Suakin Street, Pymble, New South Wales 2073, Australia under the provisions of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure and accorded AGAL Accession No. NM99/04566.

For the purposes of nomenclature, plasmid pLcUSP contains the cDNA encoding the EcR partner protein (USP polypeptide) subunit of the *Lucilia cuprina* ecdysone receptor or the USP polypeptide subunit of the *L. cuprina* juvenile hormone receptor. This plasmid was deposited on 1 July, 1999 with the Australian Government Analytical Laboratories at 1 Suakin Street, Pymble, New South Wales 2073, Australia under the provisions of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure and accorded AGAL Accession No. NM99/04565.

For the purposes of nomenclature, plasmid pMpEcR contains the cDNA encoding the EcR

polypeptide subunit of the *Myzus persicae* ecdysone receptor. This plasmid was deposited on 1 July, 1999 with the Australian Government Analytical Laboratories at 1 Suakin Street, Pymble, New South Wales 2073, Australia under the provisions of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure and accorded AGAL Accession No. NM99/04567.

For the purposes of nomenclature, plasmid pMpUSP contains a first cDNA encoding the EcR partner protein (USP polypeptide) subunit of the *Myzus persicae* ecdysone receptor or the USP polypeptide subunit of the *M. persicae* juvenile hormone receptor. This plasmid was deposited on 1 July, 1999 with the Australian Government Analytical Laboratories at 1 Suakin Street, Pymble, New South Wales 2073, Australia under the provisions of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure and accorded AGAL Accession No. NM99/04568.

For the purposes of nomenclature, plasmid pMpUSP2 contains a second cDNA encoding the EcR partner protein (USP polypeptide) subunit of the *Myzus persicae* ecdysone receptor or the USP polypeptide subunit of the *M. persicae* juvenile hormone receptor. This plasmid was deposited on 21 June, 2000 with the Australian Government Analytical Laboratories at 1 Suakin Street, Pymble, New South Wales 2073, Australia under the provisions of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure and accorded AGAL Accession No. NM00/12581.

For the purposes of nomenclature, plasmid pBtUSP contains the cDNA encoding the EcR partner protein (USP polypeptide) subunit of the *Bemesia tabaci* ecdysone receptor or the USP polypeptide subunit of the *B. tabaci* juvenile hormone receptor. This plasmid was deposited on 21 June, 2000 with the Australian Government Analytical Laboratories at 1 Suakin Street, Pymble, New South Wales 2073, Australia under the provisions of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure and accorded AGAL Accession No. NM00/12580.

The deposits referred to herein will be maintained under the Budapest Treaty on the

International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure. These deposits are provided merely for the purposes of exemplification and are not an admission that a deposit is required under 35USC §112. A license may be required to make, use or sell the deposited materials or a polypeptide encoded by a cDNA thereof and no such license is hereby granted. It is to be understood however, that the deposits will become publicly available upon the grant of a patent pertaining to the instant disclosure in so far as that patent relates to the deposits referred to herein.

Preferably, the percentage similarity to any one of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 38, SEQ ID NO: 40, or SEQ ID NO: 42, or to a polypeptide encoded by a cDNA contained in a plasmid selected from the group consisting of AGAL Accession Nos. NM99/04565, NM99/04566, NM99/04567, NM99/04568, NM00/12580, and NM00/12581, is at least about 60%, more preferably at least about 80%, even more preferably at least about 90%.

In determining whether or not two amino acid sequences fall within these percentage limits, those skilled in the art will be aware that it is necessary to conduct a side-by-side comparison or multiple alignment of sequences. In such comparisons or alignments, differences will arise in the positioning of non-identical residues, depending upon the algorithm used to perform the alignment. In the present context, reference to a percentage identity or similarity between two or more amino acid sequences shall be taken to refer to the number of identical and similar residues respectively, between said sequences as determined using any standard algorithm known to those skilled in the art. For example, amino acid sequence identities or similarities may be calculated using the GAP programme and/or aligned using the PILEUP programme of the Computer Genetics Group, Inc., University Research Park, Madison, Wisconsin, United States of America (Devereaux *et al*, 1984). The GAP programme utilizes the algorithm of Needleman and Wunsch (1970) to maximise the number of identical/similar residues and to minimise the number and length of sequence gaps in the alignment. Alternatively or in addition, wherein more than two amino acid sequences are being compared, the ClustalW programme of Thompson *et al* (1994) is used.

In an alternative embodiment, the isolated nucleic acid molecule of the invention encodes or is complementary to an isolated nucleic acid molecule which encodes a steroid receptor polypeptide or a fragment thereof, or a partner protein (USP) or a fragment thereof, which at least comprises an amino acid sequence which is substantially identical to a sequence selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 38, SEQ ID NO: 40, and SEQ ID NO: 42, or substantially identical to the amino acid sequence of a polypeptide encoded by cDNA contained in a plasmid selected from the group consisting of AGAL Accession Nos. NM99/04565, NM99/04566, NM99/04567, NM99/04568, NM00/12580, and NM00/12581.

As used herein, the term "substantially identical" or similar term shall be taken to include any sequence which is at least about 95% identical and preferably at least 99% or 100% identical to a stated nucleotide sequence or amino acid sequence, including any homologue, analogue or derivative of said stated nucleotide sequence or amino acid sequence.

Those skilled in the art will be aware that variants of the nucleotide sequences set forth in any one of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 37, SEQ ID NO: 39, or SEQ ID NO: 41, or variants of the cDNAs contained in any one of the deposited plasmids, which variants encode EcR polypeptides of insect steroid receptors or fragments thereof or EcR partner proteins (USP polypeptides) or fragments thereof, or USP polypeptides of insect juvenile hormone receptors, may be isolated by hybridization under low stringency conditions as exemplified herein.

Such variants include any genomic sequences, cDNA sequences mRNA or other isolated nucleic acid molecules derived from the nucleic acid molecules exemplified herein by the Sequence Listing. Additional variants are not excluded.

In a particularly preferred embodiment of the invention, the variant nucleotide sequences encode a fragment of the EcR polypeptide of the insect steroid receptor or a fragment of the

EcR partner protein (USP polypeptide) of the insect steroid receptor or a fragment of the USP polypeptide of the insect juvenile hormone receptor.

Preferred fragments of the subject polypeptides include one or more regions or domains which are involved in the interaction or association between the monomeric polypeptide subunits of a multimeric receptor and/or which are involved in the interaction or association between (i) a cognate steroid or receptor ligand or *cis*-acting DNA sequence; and (ii) said monomeric polypeptide subunits or the receptor *per se*. In a particularly preferred embodiment, the fragments comprise the DNA-binding domain, linker domain (domain D) or a part thereof, or ligand-binding domain (eg. hormone-binding domain) of a steroid receptor polypeptide or juvenile hormone receptor polypeptide or receptor holoenzyme. As exemplified herein, wherein biological activity of the *L. cuprina* ecdysone receptor is required, it is preferably to include at least a ligand-binding region comprising the ligand-binding domain and at least a part of the linker domain of the EcR polypeptide subunit, optionally in association with a ligand-binding region comprising at least the ligand-binding domain and at least a part of the linker domain of the EcR partner protein (USP polypeptide) subunit of said receptor. Additional fragments are not excluded.

Homologues, analogues and derivatives of the nucleotide sequences exemplified herein may be isolated by hybridising same under at least low stringency conditions and preferably under at least medium stringency conditions, to a nucleotide sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 37, SEQ ID NO: 39, and SEQ ID NO: 41, or a complementary strand of any one of said sequences, or to a cDNA contained in any one or more of the deposited plasmids. More preferably, the isolated nucleic acid molecule according to this aspect of the invention is capable of hybridising under at least high stringency conditions to a nucleic acid molecule selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 37, SEQ ID NO: 39, and SEQ ID NO: 41, or to a complementary strand of any one of said sequences, or to the cDNAs contained in any one or more of the deposited plasmids.

For the purposes of defining the level of stringency, a low stringency is defined herein as being a hybridisation or a wash carried out in 6xSSC buffer, 0.1% (w/v) SDS at 28°C or alternatively, as exemplified herein. Generally, the stringency is increased by reducing the concentration of SSC buffer, and/or increasing the concentration of SDS and/or increasing the temperature of the hybridisation and/or wash. A medium stringency comprises a hybridisation and/or a wash carried out in 0.2xSSC-2xSSC buffer, 0.1% (w/v) SDS at 42°C to 65°C, while a high stringency comprises a hybridisation and/or a wash carried out in 0.1xSSC-0.2xSSC buffer, 0.1% (w/v) SDS at a temperature of at least 55°C. Conditions for hybridisations and washes are well understood by one normally skilled in the art. For the purposes of further clarification only, reference to the parameters affecting hybridisation between nucleic acid molecules is found in Ausubel *et al.* (1992), which is herein incorporated by reference.

In an even more preferred embodiment of the invention, a hybridising nucleic acid molecule further comprises a sequence of nucleotides which is at least 40% identical to at least 10 contiguous nucleotides, preferably at least 50 contiguous nucleotides and more preferably at least 100 contiguous nucleotides, derived from a nucleotide sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 37, SEQ ID NO: 39, and SEQ ID NO: 41, or a complementary strand of any one of said sequences, or a nucleotide sequence of a cDNA contained in any one or more of the deposited plasmids referred to herein.

In determining whether or not two nucleotide sequences fall within these percentage limits, those skilled in the art will be aware that it is necessary to conduct a side-by-side comparison or multiple alignment of sequences. In such comparisons or alignments, differences may arise in the positioning of non-identical residues, depending upon the algorithm used to perform the alignment. In the present context, reference to a percentage identity between two or more nucleotide sequences shall be taken to refer to the number of identical residues between said sequences as determined using any standard algorithm known to those skilled in the art. For example, nucleotide sequences may be aligned and their identity calculated using the BESTFIT programme or other appropriate programme of the Computer Genetics Group, Inc., University

Research Park, Madison, Wisconsin, United States of America (Devereaux *et al*, 1984).

In an alternative embodiment, nucleotide sequences encoding EcR polypeptide subunits of insect steroid receptors or fragments thereof or EcR partner proteins (USP polypeptides) of insect steroid receptor or fragments thereof, or USP polypeptides of insect juvenile hormone receptor polypeptides, are amplified in the polymerase chain reaction. According to this embodiment, one or two or more nucleic acid "primer molecules" derived from a nucleotide sequence exemplified herein, such as, for example, a sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 37, SEQ ID NO: 39, and SEQ ID NO: 41, or a complementary sequence to any one of said sequences, or a sequence from a cDNA contained in any one or more of the deposited plasmids referred to herein, are annealed or hybridized to a nucleic acid "template molecule" which at least comprises a nucleotide sequence encoding a related genetic sequence or a functional part thereof, and nucleic acid molecule copies of the template molecule are amplified enzymatically using a thermostable DNA polymerase enzyme, such as *Taq*I polymerase or *Pfu* polymerase, amongst others.

More particularly, one of the primer molecules comprises contiguous nucleotides derived from a sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 37, SEQ ID NO: 39, and SEQ ID NO: 41, or from a cDNA contained in any one or more of the deposited plasmids referred to herein; and another of said primers comprises contiguous nucleotides complementary to a sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 37, SEQ ID NO: 39, and SEQ ID NO: 41, or alternatively, from a cDNA contained in any one or more of the deposited plasmids referred to herein, subject to the proviso that the first and second primers are not complementary to each other.

In a preferred embodiment, each nucleic acid primer molecule is at least 10 nucleotides in

length, more preferably at least 20 nucleotides in length, even more preferably at least 30 nucleotides in length, still more preferably at least 40 nucleotides in length and even still more preferably at least 50 nucleotides in length.

Furthermore, the nucleic acid primer molecules consists of a combination of any of the nucleotides adenine, cytidine, guanine, thymidine, or inosine, or functional analogues or derivatives thereof which are at least capable of being incorporated into a polynucleotide molecule without having an inhibitory effect on the hybridisation of said primer to the template molecule in the environment in which it is used.

Furthermore, one or both of the nucleic acid primer molecules may be contained in an aqueous mixture of other nucleic acid primer molecules, for example a mixture of degenerate primer sequences which vary from each other by one or more nucleotide substitutions or deletions. Alternatively, one or both of the nucleic acid primer molecules may be in a substantially pure form.

In a particularly preferred embodiment exemplified herein, two primer nucleotide sequences are used to amplify related sequences, said primers comprising a nucleotide sequence selected from the group consisting of SEQ ID NOs: 23 to 30 inclusive. Even more preferably, the primers are used in a primer combination selected from the group consisting of (i)SEQ ID NO: 23 and SEQ ID NO: 24; (ii) SEQ ID NO: 25 and SEQ ID NO: 26; (iii) SEQ ID NO: 27 and SEQ ID NO: 28; and (iv) SEQ ID NO: 31 and SEQ ID NO: 32.

The nucleic acid template molecule may be in a recombinant form, in a virus particle, insect cell, bacteriophage particle, yeast cell, animal cell, or a plant cell. Preferably, the nucleic acid template molecule is derived from an insect species.

Those skilled in the art will be aware that there are many known variations of the basic polymerase chain reaction procedure. Such variations are discussed, for example, in McPherson *et al* (1991). The present invention extends to the use of all such variations in the isolation of variant insect steroid receptor-encoding genes or fragments thereof, or variant

partner protein-encoding genes or fragments thereof to those exemplified herein.

The isolated nucleic acid molecule of the present invention, including those sequences exemplified herein and any variants thereof, may be cloned into a plasmid or bacteriophage molecule, for example to facilitate the preparation of primer molecules or hybridisation probes or for the production of recombinant gene products. Methods for the production of such recombinant plasmids, cosmids, bacteriophage molecules or other recombinant molecules are well-known to those of ordinary skill in the art and can be accomplished without undue experimentation. Accordingly, the invention further extends to any recombinant plasmid, bacteriophage, cosmid or other recombinant molecule comprising a nucleotide sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 24, SEQ ID NO: 25, SEQ ID NO: 26, SEQ ID NO: 27, SEQ ID NO: 28, SEQ ID NO: 31, SEQ ID NO: 32, SEQ ID NO: 37, SEQ ID NO: 39, and SEQ ID NO: 41, or a complementary sequence to any one of said sequences, or a homologue, analogue or derivative of any one of said sequences or complements, or a cDNA contained in any one or more of the deposited plasmids referred to herein.

The nucleic acid molecule of the present invention is also useful for developing genetic constructs which comprise and preferably, express, the EcR polypeptide subunit of the insect steroid receptor and/or the EcR partner protein (USP polypeptide) of the steroid receptor or the USP polypeptide of the juvenile hormone receptor, thereby providing for the production of the recombinant polypeptides in isolated cells or transformed tissues.

Accordingly, a further aspect of the present invention provides a genetic construct comprising the subject isolated nucleic acid molecule encoding the insect steroid receptor polypeptide or a juvenile hormone receptor polypeptide, operably linked to a promoter sequence. Preferably, the subject nucleic acid molecule is in an expressible format, such that it is possible to produce a recombinant polypeptide therefrom.

Reference herein to a "promoter" is to be taken in its broadest context and includes the transcriptional regulatory sequences of a classical genomic gene, including the TATA box which is required for accurate transcription initiation in a eukaryotic cell, with or without a CCAAT box sequence or alternatively, the Pribnow box required for accurate expression in prokaryotic cells.

Promoters may be cell, tissue, organ or system specific, or may be non-specific. Using specific promoters, the expression of a bioactive agent or other polypeptide encoded by a structural gene to which the promoter is operably connected may be targeted to a desired cellular site. For example, in transgenic animals such as sheep, it can be envisaged that cells of the transgenic animal may contain a gene encoding a steroid receptor, preferably a steroid receptor linked to an epidermal specific promoter and a separate gene encoding, for example, epidermal growth factor (EGF) which is functionally linked to one or more insect hormone response elements and may or may not also be linked to epidermal specific promoter elements. On administration of the appropriate insect steroid hormone to the transgenic animal, the activated complex between the insect steroid receptor and insect steroid may bind to the one or more insect steroid hormone response element thereby inducing EGF production solely in epidermal cells which may give rise to defleecing. It is to be understood that this aspect of the invention is independent of the degree of thermostability of the insect steroid receptor The same principal applies to expression of any bioactive molecule or reporter molecule in a specific cell type which is regulated by a transactivating complex between a steroid receptor complex and an appropriate insect steroid.

In the present context, the term "promoter" is also used to describe a synthetic or fusion molecule, or derivative which confers, activates or enhances expression in a cell in response to an external stimulus. Accordingly, the promoter may include further regulatory elements (i.e. upstream activating sequences, enhancers and silencers) which alter gene expression in response to developmental and/or external stimuli, or in a tissue-specific manner. Preferred promoters may contain copies of one or more specific regulatory elements, in particular steroid responsive elements (SREs) or hormone-responsive elements (HREs), to further enhance expression and/or to alter the spatial expression and/or temporal expression pattern.

Reference herein to the term "steroid response element" shall be taken to refers to one or more *cis*-acting nucleotide sequences present in a naturally-occurring or synthetic or recombinant gene the expression of which is regulated by an insect steroid, such as an ecdysteroid, for example ecdysone or ponasterone A, wherein said regulation of expression results from an direct or indirect interaction between a steroid receptor and said *cis*-acting nucleotide sequence response element. Exemplary insect steroid hormone response elements include the ecdysone response element hsp27 (EcRE) and any other nucleotide sequence which is capable of binding ecdysteroid receptors or polypeptide subunits thereof or fragments or analogies thereof (such as associated with E75, E74 or other *Drosophila* early genes), as described for example by Riddihough and Pelham (1987).

For example, an SRE or a plurality of such elements may be operably linked to a promoter such as the polyhedron promoter, p10 promoter, MMTV promoter or SV40 promoter, to make transcription of a structural gene to which said promoter is operably connected responsive to the presence of a steroid bound to the insect receptor (which may act as a transcription factor). One or more insect SREs may be located within a promoter, and may replace sequences within a selected promoter which confer responsiveness to hormones or other agents which regulate promoter activity. Where response elements are different they may lead to preferential binding of different insect steroids or analogues thereof such that a promoter may be differentially regulated.

Particularly preferred SREs according to this embodiment include, but are not limited to, the hsp27 ecdysone response element described by Riddihough and Pelham (1987) or the 13 base-pair palindromic core contained therein.

A promoter is usually, but not necessarily, positioned upstream or 5′, of a structural gene, the expression of which it regulates. Furthermore, the regulatory elements comprising a promoter are usually positioned within 2 kb of the start site of transcription of the gene.

Placing a gene or isolated nucleic acid molecule operably under the control of a promoter sequence means positioning said gene or isolated nucleic acid molecule such that its

expression is controlled by the promoter sequence. Promoters are generally positioned 5′ (upstream) to the genes that they control. In the construction of heterologous promoter/structural gene combinations it is generally preferred to position the promoter at a distance from the gene transcription start site that is approximately the same as the distance between that promoter and the gene it controls in its natural setting, i.e., the gene from which the promoter is derived. As is known in the art, some variation in this distance can be accommodated without loss of promoter function. Similarly, the preferred positioning of a regulatory sequence element with respect to a heterologous gene to be placed under its control is defined by the positioning of the element in its natural setting, i.e., the genes from which it is derived. Again, as is known in the art, some variation in this distance can also occur.

Those skilled in the art will recognise that the choice of promoter will depend upon the nature of the cell being transformed and when expression is required. Furthermore, it is well-known in the art that the promoter sequence used in the expression vector will also vary depending upon the level of expression required and whether expression is intended to be constitutive or regulated.

For expression in eukaryotic cells, the genetic construct generally comprises, in addition to the nucleic acid molecule of the invention, a promoter and optionally other regulatory sequences designed to facilitate expression of said nucleic acid molecule. The promoter may be derived from a genomic clone which normally encodes the expressed protein or alternatively, it may be a heterologous promoter derived from another genetic source. Promoter sequences suitable for expression of genes in eukaryotic cells are well-known in the art.

Suitable promoters for use in eukaryotic expression vectors include those capable of regulating expression in mammalian cells, insect cells such as Sf9 or Sf21. (*Spodoptera frugiperda*) cells, yeast cells and plant cells. Preferred promoters for expression in eukaryotic cells include the p10 promoter, MMTV promoter, polyhedron promoter, the SV40 early promoter and the cytomegalovirus (CMV- IE) promoter, promoters derived from immunoglobulin-producing cells (see, United States Patent No 4,663,281), polyoma virus promoters, and the LTR from various retroviruses (such as murine leukemia virus, murine or Rous sarcoma virus and HIV), amongst

others (See, *Enhancers and Eukaryotic Gene Expression*, Cold Spring Harbor Press, New York, 1983, which is incorporated herein by reference). Examples of other expression control sequences are enhancers or promoters derived from viruses, such as SV40, Adenovirus, Bovine Papilloma Virus, and the like.

Wherein the expression vector is intended for the production of recombinant protein, the promoter is further selected such that it is capable of regulating expression in a cell which is capable of performing any post-translational modification to the polypeptide which may be required for the subject recombinant polypeptide to be functional, such as N-linked glycosylation. Cells suitable for such purposes may be readily determined by those skilled in the art. By way of exemplification, Chinese hamster ovary (CHO) cells may be employed to carry out the N-terminal glycosylation and signal sequence cleavage of a recombinant polypeptide produced therein. Alternatively, a baculovirus expression vector such as the pFastBac vector supplied by GibcoBRL may be used to express recombinant polypeptides in Sf9 (*Spodoptera frugiperda*) cells, following standard protocols.

Numerous expression vectors suitable for the present purpose have been described and are readily available. The expression vector may be based upon the pcDNA3 vector distributed by Medos Company Pty Ltd, Victoria, Australia, which comprises the CMV promoter and BGH terminator sequences for regulating expression of the recombinant polypeptide of the invention in a eukaryotic cell, when isolated nucleic acid sequences encoding same are inserted, in the sense orientation relative to the CMV promoter, into the multiple cloning site of said vector. Alternatively, the SG5 expression vector of Greene *et al.* (1988), supplied by Stratagene, or the pQE series of vectors supplied by Qiagen are particularly useful for such purposes, as exemplified herein.

Examples of eukaryotic cells contemplated herein to be suitable for expression include mammalian, yeast, insect, plant cells or cell lines such as COS, VERO, HeLa, mouse C127, Chinese hamster ovary (CHO), WI-38, baby hamster kidney (BHK), MDCK, sf21 (insect) or Sf9 (insect) cell lines. Such cell lines are readily available to those skilled in the art.

The prerequisite for expression in prokaryotic cells such as *Escherichia coli* is the use of a strong promoter with an effective ribosome binding site. Typical promoters suitable for expression in bacterial cells such as *E. coli* include, but are not limited to, the *lacz* promoter, temperature-sensitive λ_L or λ_R promoters, T7 promoter or the IPTG-inducible *tac* promoter. A number of other vector systems for expressing the nucleic acid molecule of the invention in *E.coli* are well-known in the art and are described for example in Ausubel *et al* (1992).

Numerous vectors having suitable promoter sequences for expression in bacteria have been described, such as for example, pKC30 (λ_L:Shimatake and Rosenberg, 1981), pKK173-3 (*tac*: Amann and Brosius, 1985), pET-3 (T7: Studier and Moffat, 1986) or the pQE series of expression vectors (Qiagen, CA), amongst others.

Suitable prokaryotic cells include corynebacterium, salmonella, *Escherichia coli, Bacillus* sp. and *Pseudomonas* sp, amongst others. Bacterial strains which are suitable for the present purpose are well-known in the relevant art (Ausubel *et al*, 1992).

The genetic constructs described herein may further comprise genetic sequences corresponding to a bacterial origin of replication and/or a selectable marker gene such as an antibiotic-resistance gene, suitable for the maintenance and replication of said genetic construct in a prokaryotic or eukaryotic cell, tissue or organism. Such sequences are well-known in the art.

Selectable marker genes include genes which when expressed are capable of conferring resistance on a cell to a compound which would, absent expression of said selectable marker gene, prevent or slow cell proliferation or result in cell death. Preferred selectable marker genes contemplated herein include, but are not limited to antibiotic-resistance genes such as those conferring resistance to ampicillin, Claforan, gentamycin, G-418, hygromycin, rifampicin, kanamycin, neomycin, spectinomycin, tetracycline or a derivative or related compound thereof or any other compound which may be toxic to a cell.

The origin of replication or a selectable marker gene will be spatially-separated from those

genetic sequences which encode the recombinant receptor polypeptide or fusion polypeptide comprising same.

Preferably, the genetic constructs of the invention, including any expression vectors, are capable of introduction into, and expression in, an *in vitro* cell culture, or for introduction into, with or without integration into the genome of a cultured cell, cell line or transgenic animal. In a particularly preferred embodiment, the expression vector is selected from the group consisting of: pLcEcR (AGAL Accession No. NM99/04566); pLcUSP (AGAL Accession No. NM99/04565); pMpEcR (AGAL Accession No. NM99/04567); pMpUSP (AGAL Accession No. NM99/04568.); pMpUSP2 (AGAL Accession No. NM00/12581); and pBtUSP (AGAL Accession No. NM00/12580).

A further aspect of the invention provides a cell comprising the subject isolated nucleic acid molecule which encodes a steroid receptor polypeptide or a juvenile hormone receptor polypeptide.

As used herein, the word "cell" shall be taken to refer to a single cell, or a cell lysate, or a tissue, organ or whole organism comprising same, including a tissue, organ or whole organism comprising a clonal group of cells or a heterogenous mixture of cell types, which may be a prokaryotic or eukaryotic cell as described *supra*.

In a preferred embodiment, the cell of the present invention expresses the isolated or recombinant polypeptide encoded by the nucleic acid molecule.

In a preferred embodiment, the cell expresses a steroid receptor polypeptide or a fragment thereof which receptor is capable of binding to an insect steroid or analogue thereof or a candidate insecticidally active agent to form an activated complex, and comprises a nucleic acid sequence encoding a bioactive molecule or a reporter molecule operably linked to one or more insect steroid response elements which on binding of the said activated complex promotes transcription of the nucleic acid sequence, wherein said cell on exposure to insect steroid or an analogue thereof, regulates expression of said bioactive molecule or allows detection of said

reporter molecule.

To produce the cells of the invention, host cells are transfected or co-transfected or transformed with nucleotide sequences containing the DNA segments of interest (for example, the insect steroid receptor gene, the recombinant steroid response elements, or both) by well-known methods, which vary depending on the type of cellular host. For example, calcium chloride transfection is commonly utilized for prokaryotic cells, whereas lipofection or calcium phosphate treatment are often used for other cellular hosts. See, generally, Sambrook *et al*, (1989); Ausubel *et al*, (1992); and Potrykus (1990). Other transformation techniques include electroporation, DEAE-dextran, microprojectile bombardment, lipofection, microinjection, and others.

As used herein, the term "transformed cell" is meant to also include the progeny of a transformed cell.

In a further aspect of this invention, there is provided an animal (such as a mammal or insect), microorganism, plant or aquatic organism, containing one or more cells as mentioned above. Reference to plants, microorganisms and aquatic organisms includes any such organisms.

In this embodiment of the invention, it is to be appreciated that administration of an insect steroid or an analogue thereof to an organism will induce expression of the desired bioactive molecule, such as a polypeptide, with attendant advantages. For example, an induced protein may have a therapeutic effect ameliorating a disease state or preventing susceptibility to disease or may modify in some way the phenotype of an organism to produce a desired effect. In humans, for example, cell transplants (such as liver cells) may under the action of insect steroids, produce desirable hormones such as insulin, growth hormone, growth factors and the like.

A further aspect of the invention provides a recombinant or isolated polypeptide comprising a steroid receptor polypeptide or juvenile hormone receptor polypeptide derived from an insect or a bioactive derivative or analogue thereof, wherein said polypeptide:

(i) is selected from the group consisting of the EcR polypeptide of a steroid receptor, the partner protein (USP polypeptide) of a steroid receptor and the USP polypeptide of a juvenile hormone receptor; and

(ii) comprises an amino acid sequence having at least 40% identity to an amino acid sequence selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 38, SEQ ID NO: 40, and SEQ ID NO: 42;

wherein said polypeptide is substantially free of naturally-associated insect cell components.

In an alternative embodiment, the recombinant or isolated polypeptide comprising a steroid receptor polypeptide or juvenile hormone receptor polypeptide derived from an insect or a bioactive derivative or analogue thereof, wherein said polypeptide:

(i) is selected from the group consisting of the EcR polypeptide of a steroid receptor, the partner protein (USP polypeptide) of a steroid receptor and the USP polypeptide of a juvenile hormone receptor; and

(ii) comprises an amino acid sequence having at least 40% identity to an amino acid sequence encoded by cDNA present in a plasmid selected from the group consisting of AGAL Accession Nos. NM99/04565, NM99/04566, NM99/04567, NM99/04568, NM00/12580, and NM00/12581;

wherein said polypeptide is substantially free of naturally-associated insect cell components.

Reference herein to "substantially free of naturally associated insect cell components" refers to at least 80% purity, preferably more than 90% purity, and more preferably more than 95% purity. Normally, purity is measured on a polyacrylamide gel with homogeneity determined by staining of protein bonds. Alternatively, high resolution may be necessary using HPLC or similar means. For most purposes, a simple chromatography column or polyacrylamide gel may be used to determine purity. A protein which is chemically synthesized or synthesized in a cell system different from an insect cell from which it naturally originates would be free of naturally-associated insect cell components.

The present invention clearly provides for the isolation of EcR polypeptide subunits and EcR

partner protein (USP polypeptide) subunits of ecdysteroid receptors and USP polypeptides of juvenile hormone receptors, from various organisms of the class *Insecta*, as described *supra*, in addition to protozoa and helminth sources.

Insect steroid receptors are characterized by functional ligand-binding domains, and DNA-binding domains, both of which interact to effect a change in the regulatory state of a gene operably linked to the DNA-binding site of the holoreceptor or a polypeptide or polypeptide fragment thereof. Thus, insect steroid receptors seem to be ligand-responsive transcription factors. Additionally, insect steroid receptors generally contain a DNA-binding domain (Domain C), and a ligand-binding domain (Domain E), separated and flanked by additional domains as identified by Krust *et al* (1986). The C domain preferably comprises a zinc-finger DNA-binding domain which is usually hydrophilic, having high cysteine, lysine and arginine content. The E domain preferably comprises hydrophobic amino acid residues and is further characterized by regions E1, E2 and E3. The ligand-binding domain of the members of the insect steroid receptor superfamily is typically carboxyl-proximal, relative to a DNA-binding domain (Evans,1988). The entire ligand-binding domain is typically between about 200 and 250 amino acids but is potentially shorter. This domain has the subregions of high homology, designated the E1, E2 and E3 regions - which may be collectively referred to as the "E region". Amino acid residues proximal to the C domain comprise a region initially defined as separate A and B domains. Region D separates the more conserved domains C and E. Region D typically has a hydrophilic region whose predicted secondary structure is rich in turns and coils. The F region is carboxy promixal to the E region (see, Krust *et al*, *supra*).

The receptor polypeptides of the present invention exhibit at least a ligand-binding domain, as characterized by sequence homology to regions E1, E2 and E3. The ligand-binding domains of the present invention are typically characterized by having significant homology in sequence and structure to these three regions. Fragments of insect steroid receptors and partner proteins capable of binding insect steroids, and candidate insecticidally active compounds comprise an E-region or a sufficient portion of the E-region to allow binding.

Preferably, the recombinant or isolated EcR polypeptide subunit of the insect steroid receptor

or EcR partner protein (USP polypeptide) subunit of the steroid receptor or USP polypeptide of the juvenile hormone receptor as described herein is thermostable.

By "thermostable" is meant that a stated integer does not exhibit reduced activity at bacterial, plant or animal physiological temperatures above about 28°C or above about 30°C. The thermostability of insect steroid hormone receptors also refers to the capacity of such receptors to bind to ligand-binding domains or regions or to transactivate genes linked to insect steroid hormone response elements at bacterial, plant or animal physiological temperatures above about 28°C or above about 30°C.

The present invention clearly extends to variants of said polypeptides, as described *supra*. The polypeptide may be substantially free of naturally associated insect cell components, or may be in combination with a partner protein which associates with the insect steroid receptor so as to confer enhanced affinity for insect steroid response elements, enhanced affinity for insect steroids or analogues thereof. For Example, the amino acid sequences exemplified herein may be varied by the deletion, substitution or insertion of one or more amino acids.

In one embodiment, amino acids of a polypeptide exemplified herein may be replaced by other amino acids having similar properties, for example hydrophobicity, hydrophilicity, hydrophobic moment, charge or antigenicity, and so on.

Substitutions encompass amino acid alterations in which an amino acid of the base polypeptide is replaced with a different naturally-occurring or a non-conventional amino acid residue. Such substitutions may be classified as "conservative", in which case an amino acid residue contained in the base polypeptide is replaced with another naturally-occurring amino acid of similar character, for example Gly↔Ala, Val↔Ile↔Leu, Asp↔Glu, Lys↔Arg, Asn↔Gln or Phe↔Trp↔Tyr.

Substitutions encompassed by the present invention may also be "non-conservative", in which an amino acid residue which is present in the base polypeptide is substituted with an amino acid having different properties, such as a naturally-occurring amino acid from a different group

(eg. substituted a charged or hydrophobic amino acid with alanine), or alternatively, in which a naturally-occurring amino acid is substituted with a non-conventional amino acid.

Those skilled in the art will be aware that several means are available for producing variants of the exemplified EcR polypeptide subunit of the insect steroid receptor or EcR partner protein (USP polypeptide) subunit of the steroid receptor or USP polypeptide of the juvenile hormone receptor, when provided with the nucleotide sequence of the nucleic acid molecule which encodes said polypeptide, for example site-directed mutagenesis of DNA and polymerase chain reaction utilising mutagenised oligonucleotide primers, amongst others.

Such polypeptide variants which are capable of binding insect steroids clearly form part of the present invention. Assays to determine such binding may be carried out according to procedures well known in the art.

One such variant polypeptide encompassed by the present invention comprises an "in-frame" fusion polypeptide between different regions of different insect receptor polypeptides. As exemplified herein, the present inventors have discovered that, by producing synthetic genes in which various domains of a base insect steroid receptor-encoding nucleotide sequence derived from a first source are interchanged or substituted with similar sequences derived from a second source (referred to as "domain swapping"), it is possible to modify the bioactivity of the insect steroid receptor encoded therefor. For example, the biological activity of the EcR polypeptide of the *L. cuprina* or *M. persicae* ecdysone receptor exemplified herein may be modulated by replacing portions of its C-terminal or N-terminal sequences with the equivalent domains from the EcR polypeptide of the *D. melanogaster* ecdysone receptor or alternatively, by swapping regions of the EcR polypeptides of the *L. cuprina* and *M. persicae* ecdysone receptors *per se*.

As a further refinement, such changes in biological function can similarly be effected by making specific changes (e.g. addition, substitution or deletion) to only those amino-acids within each domain that are critical for determining the relevant catalytic function (eg. ligand-binding activity, DNA binding site affinity, etc), such as by site-directed mutagenesis.

According to this embodiment, there is provided a synthetic EcR polypeptide subunit of a steroid receptor, or a synthetic EcR partner protein (USP polypeptide) subunit of a steroid receptor, or a synthetic USP polypeptide of a juvenile hormone receptor, or an analogue or derive of said synthetic polypeptides, wherein said synthetic polypeptides comprise an amino acid sequence which has the following properties:

(i) it differs in amino acid sequence or exhibits different biological properties to a naturally-occurring EcR polypeptide subunit of a steroid receptor, or a naturally-occurring EcR partner protein (USP polypeptide) subunit of a steroid receptor, or a naturally-occurring USP polypeptide of a juvenile hormone receptor;

(ii) it comprises a first sequence of amino acids having at least about 40% identity to a part of an amino acid sequence selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 38, SEQ ID NO: 40, and SEQ ID NO: 42, or having at least about 40% identity to a part of an amino acid sequence encoded by any one of the deposited plasmids referred to herein, linked covalently to a second sequence of amino acids derived from an EcR polypeptide subunit of a steroid receptor, EcR partner protein (USP polypeptide) subunit of a steroid receptor, or USP polypeptide of a juvenile hormone receptor, wherein said first and second sequences are derived from different genomic sources.

Preferably, the first sequence of amino acids is derived from the EcR polypeptide subunit of a steroid receptor, more preferably from the EcR polypeptide of the *L. cuprina* or *M. persicae* ecdysone receptor, and even more preferably from the EcR polypeptide of the *L. cuprina* ecdysone receptor.

In one embodiment, the synthetic EcR polypeptide subunit of a steroid receptor, or a synthetic EcR partner protein (USP polypeptide) subunit of a steroid receptor, or a synthetic USP polypeptide of a juvenile hormone receptor comprises a fusion polypeptide in which the ligand-binding regions of an amino acid sequence selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 38, SEQ ID NO: 40, and SEQ ID NO: 42 are

replaced, in-frame, by the ligand-binding region of a different receptor polypeptide.

In a particularly preferred embodiment, 5′-end of the open reading frame of a first nucleotide sequence, encoding the N-terminal portion of the EcR polypeptide of a first ecdysteroid receptor to the end of the DNA-binding domain of said polypeptide, is fused in-frame, to the 3′-end of the open reading frame of a second nucleotide sequence, encoding the C-terminal portion of the EcR polypeptide of a second ecdysteroid receptor, from the D domain and hormone-binding domain to the carboxyl terminus.

Accordingly, the present invention extends to any variants of the insect receptor polypeptides referred to herein and genetic sequences encoding same, wherein said variants are derived from a receptor polypeptide as described herein and exhibit demonstrable ligand-binding activity, and either comprises an amino acid sequence which differs from a naturally-occurring receptor polypeptide, or exhibit biological activity.

As with other aspects of the invention, the variants described herein may be produced as recombinant polypeptides or in transgenic organisms, once the subject synthetic genes are introduced into a suitable host cell and expressed therein.

In an alternative embodiment, the recombinant receptor polypeptide of the invention is produced as an "in-frame" fusion polypeptide with a second polypeptide, for example a detectable reporter polypeptide such as β-galactosidase, β-glucuronidase, luciferase or other enzyme, or a FLAG peptide, hapten peptide such as a poly-lysine or poly-histidine or other polypeptide molecule.

By "in-frame" means that a nucleotide sequence which encodes a first polypeptide is placed (i.e. cloned or ligated) in the same open reading frame adjacent to a nucleotide sequence which encodes a second polypeptide with no intervening stop codons there between, such that when the ligated nucleic acid molecule is expressed, a single fusion polypeptide is produced which comprises a sequence of amino acids corresponding to the summation of the individual amino acid sequences of the first and second polypeptides.

In order to produce a fusion polypeptide, the nucleic acid molecule which encodes the polypeptide of the invention, or an analogue or derivative thereof, is cloned adjacent to a second nucleic acid molecule encoding the second polypeptide, optionally separated by a spacer nucleic acid molecule which encodes one or more amino acids (eg: poly-lysine or poly histidine, amongst others), such that the first coding region and the second coding region are in the same open reading frame, with no intervening stop codons between the two coding regions. When translated, the polypeptide thus produced comprises a fusion between the polypeptide products of the first and second coding regions. Wherein a spacer nucleic acid molecule is utilised in the genetic construct, it may be desirable for said spacer to at least encode an amino acid sequence which is cleavable to assist in separation of the fused polypeptide products of the first and second coding regions, for example a thrombin cleavage site.

A genetic construct which encodes a fusion polypeptide further comprises at least one start codon and one stop codon, capable of being recognised by the cell's translational machinery in which expression is intended.

Preferably, a genetic construct which encodes a fusion polypeptide may be further modified to include a genetic sequence which encodes a targeting signal placed in-frame with the coding region of the nucleotide sequence encoding the fusion polypeptide, to target the expressed recombinant polypeptide to the extracellular matrix or other cell compartment. More preferably, the genetic sequence encoding targeting signal is placed in-frame at the 5′-terminus or the 3′-terminus, but most preferably at the 5′-terminus, of the coding region of the nucleotide sequence which encodes the fusion polypeptide.

Methods for the production of a fusion polypeptide are well-known to those skilled in the art.

The recombinant EcR polypeptide subunit of the insect steroid receptor or EcR partner protein (USP polypeptide) subunit of the steroid receptor or USP polypeptide of the juvenile hormone receptor may be purified by standard techniques, such as column chromatography (using various matrices which interact with the protein products, such as ion exchange matrices,

hydrophobic matrices and the like), affinity chromatography utilizing antibodies specific for the protein or other ligands such as dyes or insect steroids which bind to the protein.

Wherein the recombinant polypeptide is expressed as a fusion polypeptide, it is also possible to purify the fusion polypeptide based upon its properties (eg size, solubility, charge etc). Alternatively, the fusion polypeptide may be purified based upon the properties of the non-receptor moiety of said fusion polypeptide, for example substrate affinity. Once purified, the fusion polypeptide may be cleaved to release the intact polypeptide of the invention.

Alternatively, proteins may be synthesized by standard protein synthetic techniques as are well known in the art.

In a preferred embodiment, the recombinant or isolated polypeptides of the invention are provided as a precipitate or crystallized by standard techniques, preferably for X-ray crystal structure determination.

The three-dimensional structure of the polypeptide of the invention or a holoreceptor comprising same or a fragment of said polypeptide or holoreceptor is particularly useful for identifying candidate insecticidal agents which mimic ligands that bind to said three-dimensional structure and/or modulate the ability of insect steroids to bind thereto and activate the receptor (see, for example, Von Itzstein *et al.,* 1993; and Bugg *et al.*, 1993).

According to this embodiment, the EcR polypeptides of the invention or ligand binding domains thereof, or their complexes with EcR partner proteins or ligand binding domains thereof, which confer enhanced affinity for insect steroid response elements or partner proteins (USP polypeptides) or ligands, are particularly useful to model the three-dimensional structure of the receptor ligand-binding region. In this manner, insecticidal compounds may be produced which bind to, or otherwise interact with, the ligand-binding region of the receptor, and preferably interfere with ligand binding. In the same way, compounds may be developed which have a potentiated interaction with the insect steroid receptor over and above that of the physiological insect steroid which binds to the receptor.

Accordingly, a still further aspect of the invention provides a method of identifying a candidate insecticidally-active agent comprising the steps of:

a) expressing a USP polypeptide of a juvenile hormone receptor or a fragment thereof which includes the ligand-binding region, optionally in association with an EcR polypeptide of a steroid receptor or ligand binding domain thereof, and optionally in association with an insect steroid or analogue thereof, so as to form a complex;

b) purifying or precipitating the complex;

c) determining the three-dimensional structure of the ligand binding domain of the complex; and

d) identifying compounds which bind to or associate with the three-dimensional structure of the ligand binding domain, wherein said compounds represent candidate insecticidally-active agents.

Standard procedures are used to determine the three dimensional structure of the receptor polypeptides of the invention, for example using X-ray crystallography and/or nuclear magnetic resonance analysis (*see*, for example, Bugg *et al.*, 1993; Von Itstein *et al.*, 1993).

Insecticidally-active agents contemplated herein include synthetic chemicals that mimic one or more ligands of the holoreceptor or its polypeptide subunit, or the ligand-binding region of said holoreceptor or subunit, thereby modulating binding of steroids to said holoreceptor or subunit. Preferred insecticidally-active agents include bisacylhydrazines, iridoid glycosides or other non-steroidal modulators of ecdysteroid receptors or insect juvenile hormone receptors. Additionally, because the EcR partner protein (USP polypeptide) subunits of insect steroid receptors, and the USP polypeptides of insect juvenile hormone receptors, bind insect juvenile hormones, a sesquiterpenoid group of ligands that regulate developmental transitions in insects (see Jones and Sharp, 1997), compounds which interfere with the binding of juvenile hormone are also candidate insecticides.

A further aspect of the present invention provides a method of identifying a modulator of insect steroid receptor-mediated gene expression or insect juvenile hormone receptor-mediated gene

expression comprising:

(i) assaying the expression of a reporter gene in the presence of a recombinant or isolated insect steroid receptor polypeptide or a juvenile hormone receptor polypeptide of the invention and a potential modulator; and

(ii) assaying the expression of a reporter gene in the presence of a recombinant or isolated insect steroid receptor polypeptide or a juvenile hormone receptor polypeptide of the invention and without said potential modulator; and

(ii) comparing expression of the reporter gene in the presence of the potential modulator to the expression of a reporter gene in the absence of the potential modulator,

wherein said reporter gene is placed operably under the control of a steroid response element (SRE) to which said insect steroid receptor binds or a promoter sequence comprising said SRE.

In the present context, a "modulator" is a compound or molecule that agonises or antagonises the binding properties and/or biological activity of a receptor polypeptide or holoreceptor. Preferred modulators according to this embodiment include those synthetic compounds that are suitable for use as insecticidally-active agents described *supra*.

The reporter gene may be any gene, the expression of which may be monitored or assayed readily. Preferably, the reporter gene is a structural gene that encodes a peptide, polypeptide or enzyme that is assayed readily by enzymic or immunological means, for example the β-galactosidase, β-glucuronidase, luciferase or chloramphenicol acetyltransferase (CAT) genes. Alternatively, the reporter gene may be a gene which encodes an immunologically-detectable protein, for example a FLAG peptide, poly-lysine peptide or poly-histidine peptide.

Standard methods are used to assay the expression of the reporter gene.

This embodiment of the invention may be applied directly to the identification of potential insecticidally-active compounds or alternatively, modified for such purposes by assaying for the binding (direct or indirect) of the recombinant or isolated insect steroid receptor polypeptide or a juvenile hormone receptor polypeptide of the invention to a steroid response element (SRE),

rather than by assaying for reporter gene expression. According to this alternative embodiment, the binding assayed in the presence or absence of a potential insecticidally-active compound is compared, wherein a difference in the level of binding indicates that the candidate compound possesses potential insecticidal activity.

In addition, substances may be screened for insecticidal activity by assessing their ability to bind, *in vivo* or *in vitro,* to the intact ecdysone receptor or alternatively, the ligand-binding regions of the EcR polypeptide subunit of the ecdysone receptor (eg. SEQ ID NO: 2 or SEQ ID NO: 10 or SEQ ID NO: 14) or the EcR partner protein (USP polypeptide) of the ecdysone receptor (eg. SEQ ID NO: 4 or SEQ ID NO: 6 or SEQ ID NO: 8 or SEQ ID NO: 16 or SEQ ID NO: 18 or SEQ ID NO: 20 or SEQ ID NO: 22 or SEQ ID NO: 38 or SEQ ID NO: 40 or SEQ ID NO: 42). Competition assays involving the native insect steroid may be employed to assess insecticidal activity.

The performance of this embodiment may, for example, involve binding the insect steroid receptor polypeptide to a support such as a plurality of polymeric pins, whereafter the polypeptide resident on the plurality of pins is brought into contact with candidate insecticidal molecules for screening. The molecules being screened may be isotopically labelled so as to permit ready detection of binding. Alternatively, reporter molecules may be utilized which bind to the insect steroid receptor candidate molecule complex. Alternatively, compounds for screening may be bound to a solid support, such as a plurality of pins which are then reacted with the thermostable insect steroid receptor or complex with a partner protein. Binding may, for example, be determined again by isotopic-labelling of the receptor, or by antibody detection or use of another reporting agent.

In an alternative embodiment, insecticidally-active agent are identified using rational drug design, by expressing a USP polypeptide of a juvenile hormone receptor or a fragment thereof which includes the ligand-binding region, optionally in association with an EcR polypeptide of a steroid receptor or ligand binding domain thereof, and optionally in association with an insect steroid or analogue thereof, so as to form a complex, determining the three-dimensional structure of the ligand binding domain of the complex, and identifying compounds which bind

to or associate with the three-dimensional structure of the ligand binding domain, wherein said compounds represent candidate insecticidally-active agents.

The methods described herein for identifying modulators of gene expression and insecticidal compounds, may be performed using prokaryotic or eukaryotic cells, cell lysates or aqueous solutions.

A further aspect of this invention accordingly relates to synthetic compounds derived from the three dimensional structure of EcR polypeptides or EcR partner protein (USP polypeptide) subunits of insect steroid receptors, or fragments thereof, or insect steroid receptors or fragments thereof, or USP polypeptides of insect juvenile hormone receptors or fragments thereof, which compounds are capable of binding to said receptors which have the effects of either inactivating the receptors (and thus acting as antagonists) or potentiating the activity of the receptor.

By "derived from" it is meant that the compounds are based on the three dimensional structure of the aforementioned proteins, that is, synthesized to bind, associate or interfere with insect steroid binding or juvenile hormone binding.

The compounds may bind strongly or irreversibly to the ligand binding site or another region of the receptor or USP and act as agonists or antagonists of insect steroids, or juvenile hormone binding, or otherwise interfere with the binding of ligand, such that ecdysteroids or juvenile hormones. Such compounds would have potent insecticidal activity given the key role of insect steroids, or juvenile hormone, in insect physiology and biochemistry. Such compounds would also possess a unique specificity.

This invention is also described with reference to the following non-limiting examples.

EXAMPLE 1

Construction of a plasmid (pSV40-EcR) expressing the EcR polypeptide subunit of the *D. melanogaster* ecdysone receptor

A 3110 base-pair *Fsp*I-*Hin*dIII fragment was excised from a cDNA encoding the EcR polypeptide subunit of the *D. melanogaster* ecdysone receptor (Koelle *et al.,*1991), the excised sequence comprising the complete 2634 base pair coding region and 214 base pairs of 5′-leader sequence and 258 base pairs of 3′- untranslated sequence. The fragment was ligated into the *Bam*HI site of the expression plasmid pSG5 (Greene *et al*, 1988) to produce the expression plasmid pSV40-EcR, wherein expression of the EcR polypeptide subunit of the *Drosophila melanogaster* ecdysone receptor is placed operably under the control of the SV40 promoter sequence.

EXAMPLE 2

Construction of the reporter plasmid p(EcRE)$_7$-CAT

The reporter plasmid p(EcRE)$_7$-CAT was constructed by insertion of multiple copies (i.e. 5 to 7 copies) of the hsp27 ecdysone response element, containing a central 13 base pair palindromic ecdysone response element (EcRE), derived from the hsp27 gene (Riddihough and Pelham, 1987) into the *Hin*dIII site of the plasmid pMMTV-CAT (Hollenberg and Evans, 1988), 93 base pairs upstream of the transcription start site of the MMTV promoter, thereby operably connecting expression of the chloramphenicol acetyltransferase structural gene to regulation by an insect receptor which binds to the hsp27 ecdysone response element.

EXAMPLE 3

Cell Culture and Transient Transfection

Chinese hamster ovary (CHO) cells were maintained in 50% (v/v) Dubbecco's modified Eagle's medium (DMEM) and 50% (v/v) Hamm F12 nutrient mixture (GIBCO) supplemented with 10% (v/v) foetal bovine serum. Transfection was carried out by the DNA-calcium phosphate co-precipitation method (Ausubel *et al*, 1992). One day before transfection with the plasmids described in Examples 1 and/or 2 , or other expression plasmids, CHO cells were plated out at 5 - 8 x 10^5 cells per 6 cm diameter culture dish in the above DMEM/F12 medium. Three

hours before the addition of the DNA-calcium phosphate co-precipitate, the cells were washed with phosphate buffered saline (PBS; Sambrook *et al.*, 1989) and cultured in fresh DMEM plus 10% (v/v) foetal bovine serum. The cells were incubated in the presence of the co-precipitate for eighteen hours before excess DNA was removed by washing with PBS. The cells were then cultured for another day in DMEM/F12 supplemented with 10% (v/v) foetal bovine serum with or without added ponasterone A (PNA), before harvesting. Cells were washed with PBS, harvested by mechanical scraping in 0.25 M Tris-HCl (pH 7.8), and disrupted by three freeze-thaw cycles.

All transfections included, in addition to expression and reporter plasmids, a β-galactosidase-expressing plasmid designated pPgK-LacZ (McBurney *et al*, 1991), which served as an internal control for the efficiency of transfection, and pUC18 DNA in an amount sufficient to produce 10 µg total DNA per culture dish.

The chloramphenicol acetyltransferase (CAT) and β-galactosidase activities encoded by the reporter genes present in the reporter plasmids were assayed as described in Sambrook *et al*, (1989). Cells that were co-transfected with p(EcRE)$_7$-CAT and pSV40-EcR clearly showed induction of CAT activity in the presence of PNA, showing 50 units of activity. Controls showed negligible activity.

We have observed that the ecdysone receptor can lead to stimulation of expression from an ecdysone responsive promoter in some cell types, for example in CHO cells, but not in CV-1 cells. Whilst not being bound by any theory or mode of action, this may reflect a cell-type specific distribution of at least one other transcription factor essential for ecdysone responsiveness. To determine cell types suitable for expressing reporter genes under the control of the steroid receptor of the present invention, the cell-type specificity of ecdysone-responsive gene expression is assayed in cell-free transcription lysates derived from several target cell lines. Additionally, by fractionating or isolating the nuclear proteins of cell lines that express the reporter genes and supplementing lysates derived from non-expressing cell lines with such nuclear protein fractions or isolated proteins, any essential auxiliary factors are defined and the genes encoding them cloned. Co-transfection of the receptor-encoding genes

with genes encoding such auxiliary factors removes limitations imposed by cell-type restricted ecdysone responsiveness.

EXAMPLE 4

Testing the Effect of temperature on transient expression

To determine whether the *D. melanogaster* ecdysone receptor polypeptide is stable at physiological temperatures above about 30°C, CHO cells were transfected as described in Example 3, with the plasmid pSV40-EcR and the reporter plasmid p(EcRE)$_7$-CAT in the presence of PNA, at 30°C and 37°C.

Briefly, CHO cells were plated out at 37°C sixteen to twenty hours before transfection. After washing away the DNA, the cells were cultured for two hours in fresh medium with or without hormone and the dishes divided into duplicate sets. One set was cultured for another day at 37°C before harvesting for CAT and β-galactosidase assays. The other set was cultured for three days at 30°C before assaying enzyme activities. Results indicated a reduction in the fold-induction of gene expression regulated by the *D. melanogaster* ecdysone receptor polypeptide at 37°C, compared to the fold-induction at 30°C, as shown in Table 1.

EXAMPLE 5

Attempted screening of an *L. cuprina* genomic DNA library to isolate genes encoding the EcR polypeptide subunit of the *L. cuprina* ecdysone receptor

A 627 bp Eco - Kpn I fragment encompassing the DNA-binding domain of the EcR polypeptide subunit of the *D. melanogaster* ecdysone receptor was isolated, radioactively labelled and used to screen a *L. cuprina* genomic library constructed in bacteriophage lambda (prepared by CSIRO, division of Entomology, Canberra, Australia). In the first round of screening, twenty-four regions of the plates showed potential positive hybridization to the *D. melanogaster* probe. However, second-round screening of these 24 first round positive plaques failed to yield any plaque giving a reproducible positive signal when hybridized to the *D. melanogaster* probe.

TABLE 1

pSV40-EcR (µg/dish)	PNA (µM)	Fold-induction of expression	
		37°C	30°C
2.5	20 100	14X 59X	35X 54X
0.5	20 100	8X 47X	26X 33X
0.1	20 100	1.6X 9.0X	25X 39X

EXAMPLE 6

Cloning and characterization of a cDNA molecule encoding the EcR polypeptide of the *L. cuprina* ecdysone receptor

Rationale for amplification primer design

The nucleotide sequences of the primers Rdna3 (SEQ ID NO: 23) and Rdna4 (SEQ ID NO: 24) were derived from the amino acid sequence conserved between the DNA-binding domains of the EcR polypeptide subunits of the *D. melanogaster* and *C. tentans* ecdysone receptors. However, amino acid sequences homologous to two other members of the steroid receptor superfamily of *D. melanogaster*, *Drosophila* hormone receptor 3 (DHR3; Koelle, *et al.,* 1991) and *Drosophila* early gene (E75; Segraves and Hogness, 1990) were excluded from the primer designs, to reduce the possibility of amplifying the *L. cuprina* homologues of genes encoding DHR3 or E75 by PCR.

Amplification primers and PCR conditions

A 105 base pair DNA fragment, encoding the DNA-binding domain of the EcR polypeptide subunit of the *L. cuprina* ecdysone receptor, was amplified from the *L. cuprina* genome by PCR, by using the following degenerate primers:

Rdna3 (32mer with EcoRI site):

5′-CGGAATTCCGCCTCTGGTTA(C/T)CA(C/T)TA(C/T)AA(C/T)GC 3′ (i.e. SEQ ID NO: 23);

and

Rdna4 (32mer with BamHI site):

5′-CGCGGATCC(G/A)CACTCCTGACACTTTCG(C/T)CTCA 3′ (i.e. SEQ ID NO: 24).

Amplification reactions employed *Taq*I DNA polymerase (Promega) and the following amplification conditions:

cycle 1:	97°C/5 minutes, 50°C hold; add polymerase 50°C/5 minutes;
cycles 2-3:	72°C/3 minutes, 94°C/1 minute, 50°C/1 minute;
cycles 4-43:	72°C/3 minutes, 94°C//1 minute, 55°C/1 minute;
cycle 44:	72°C/10 minutes.

To facilitate cloning of the amplified fragments for use as hybridisation probes, the 5' end of primer Rdna3 contained an *Eco*RI site and the 5' end of primer Rdna4 contained a *Bam*HI site. The amplified *L. cuprina* gene fragments were cloned into *p*Bluescript SK+, following digestion using the enzymes *Eco*RI and *Bam*HI, purification of the digested DNA by agarose gel electrophoresis and electro elution of the product band.

Hybridisation probe preparation

For probe preparation, the insert was cut out of the *p*Bluescript SK+ vector using *Eco*R1 and *Bam*HI, and ^{32}P-labelled using the GIGAprime DNA Labelling Kit (Bresatec Limited, Adelaide, Australia) essentially according to the manufacturer's instructions, except that random primers were replaced with the specific primers Rdna3 and Rdna4 (see above). Unincorporated label was removed by size exclusion chromatography over Biogel-P60 (Biorad Ltd, Sydney, Australia). The probe was used at 10^6 cpm/ml in hybridizations.

Construction and screening of *L. cuprina* cDNA libraries

Two independent *L. cuprina* cDNA libraries derived from late third instar *L. cuprina* larvae were prepared by random priming and oligo-dT priming respectively, and cloned into the *Eco*RI site of the *Lambda/ZapII* vector (Stratagene). The primary libraries generated were subsequently amplified according to the manufacturer's instructions, using standard protocols.

Both cDNA libraries generated are superior to existing *L. cuprina* libraries in terms of their phage titre (i.e. pfu/ml) and insert sizes (0.5 - 4 kbp in both cases). In particular, the primary oligo-dT primed library comprised 4.7 x 10^6 pfu, whilst the amplified oligo-dT primed library comprised 7.5 x 10^{10} pfu/ml; the primary random-primed library comprised 1.3 x 10^6 pfu, whilst the amplified random-primed library comprised 3.4 x 10^{10} pfu/ml.

The prepared cDNA libraries were screened by lifting 500,000 plaques from each library in duplicate on to Hybond N membranes (Amersham) and hybridizing same under low stringency conditions to the ^{32}P-labelled amplification product produced using the primers Rdna3 and Rdna4 (see above). In particular, hybridisations were performed for twenty four hours at 37°C in a hybridisation solution comprising 42% (w/v) formamide; 5 x SSPE solution; 5 x Denhardt's solution; and 0.1% (w/v) sodium dodecyl sulphate, as described essentially by Ausubel *et al,* (1992) or Sambrook *et al.* (1989). The membranes were then washed at 37°C in 2XSSC solution containing 0.1% (w/v) sodium dodecyl sulphate. Following washing, positive plaques were detected by autoradiography, using XOMAT-AR film (Kodak) for two to three days, at -70°C.

Two positive-hybridising plaques were obtained from screening of the random-primed library (containing cDNA inserts comprising 561 base pairs and 1600 base pairs in length, respectively), and one positive-hybridising plaque was obtained from the screening of the oligo-dT primed library (containing a cDNA insert comprising approximately 3400 base pairs in length). *p*Bluescript phagemids containing cDNA inserts were excised *in vivo* from these positive plaques using the Exassist Helper Phage system (Stratagene).

The nucleotide sequences of the isolated cDNA clones were obtained using the USB Sequenase Version 2.5 Kit. Sequence data obtained indicated that the 561 bp and 1600 bp cDNAs encode amino acid sequences comprising the important DNA-binding domain and the hormone-binding domain of the EcR polypeptide subunit of the *L. cuprina* ecdysone receptor, whilst the 3400 bp cDNA comprises an entire 2274 bp open reading frame encoding the EcR polypeptide subunit of the *L. cuprina* ecdysone receptor. Accordingly, the 3400 bp cDNA is a full-length cDNA clone. The nucleotide sequence of the open reading frame and 3′-

untranslated region is set forth herein as SEQ ID NO: 1. The derived amino acid sequence of the EcR polypeptide subunit of the *L. cuprina* ecdysone receptor encoded by this open reading frame is set out in SEQ ID NO: 2.

EXAMPLE 7

First attempt at cloning and characterization of a cDNA molecule encoding the EcR polypeptide of the *M. persicae* ecdysone receptor

Direct screening of a *M. persicae* cDNA library was not effective in isolating a full-length cDNA encoding the EcR polypeptide of the *M. persicae* ecdysone receptor.

DNA encoding the DNA-binding domain of the EcR polypeptide of the *M. persicae* ecdysone receptor was isolated successfully, by amplification as described in Example 6 for the amplification of the homologous *L. cuprina* fragment. The amplified DNA was cloned into *p*Bluescript SK+ and the nucleotide sequence of the cloned insert was obtained using the USB Sequenase version 2.0 Kit, as described in Example 6.

Based upon the nucleotide sequence of the amplified DNA fragment, two authentic primers were synthesized as follows:

Mdna1 (23mer): 5′- GCCTCGGGGTATCACTATAACGC -3′ (i.e. SEQ ID NO: 25); and
Mdna2 (23mer): 5′- GCACTCCTGACACTTTCGTCTCA -3′ (i.e. SEQ ID NO: 26).

<u>Hybridisation probe preparation</u>

For *M. persicae* probe preparation, the amplified 105 bp DNA insert was excised from the *p*Bluescript SK+ vector using EcoRI and BamHI, and ^{32}P-labelled using the GIGAprime DNA Labelling Kit (BresaGen Limited, Adelaide, Australia) essentially according to the manufacturer's instructions, except that random primers were replaced with the specific primers Mdna1 and Mdna2 (see above). Unincorporated label was removed by size exclusion chromatography over Biogel-P60 (Biorad Ltd, Sydney, Australia). The probe was used at 10^6 cpm/ml in hybridizations.

Construction and screening of *M. persicae* cDNA libraries:

Two independent *M. persicae* cDNA libraries derived from late third instar *M. persicae* larvae were prepared by random priming and oligo-dT priming respectively, and cloned into the *Eco*RI site of the *Lambda/ZapII* vector (Stratagene). The primary libraries generated were subsequently amplified according to the manufacturer's instructions, using standard protocols.

Both cDNA libraries generated are superior to existing *M. persicae* libraries in terms of their phage titre (i.e. pfu/ml) and insert sizes (0.5 - 4 kbp in both cases). In particular, the primary oligo-dT-primed library comprised 1×10^7 pfu, whilst the amplified oligo-dT primed library comprised 1×10^{10} pfu/ml; the primary random-primed library comprised 1×10^6 pfu, whilst the amplified random-primed library comprised 2×10^{11} pfu/ml.

Additionally, a further cDNA library was produced in the Lambda ZAP Express insertion vector (Stratagene). To produce this library, cDNA derived from late third instar *M. persicae* larvae was prepared by oligo-dT priming and cloned directionally into *Eco*RI-XhoI digested vector DNA. The primary library comprised 1×10^6 pfu, whilst the amplified oligo-dT primed library comprised 1×10^9 pfu/ml, with insert sizes in the range 0.5 - >4 kbp.

The random-primed *M. persicae* cDNA phage library was screened as described in Example 6, using the *M. persicae* hybridisation probe prepared as described above.

A single positive-hybridising plaque was isolated and sequenced according to standard procedures. The nucleotide sequence of this clone is set forth herein as SEQ ID NO: 9. This cDNA clone comprises a 585bp protein-encoding sequence which encodes the DNA-binding domain of a EcR polypeptide of a putative *M. persicae* ecdysone receptor. The amino acid sequence encoded by this partial cDNA clone is set forth herein as SEQ ID NO: 6.

EXAMPLE 8

Second attempt at cloning and characterization of a cDNA molecule encoding the EcR polypeptide of the *M. persicae* ecdysone receptor

Hybridisation probe preparation

Further hybridisation probes specific for the EcR polypeptide of the *M. persicae* ecdysone receptor were generated using PCR from the Lambda ZAPII oligo dT-primed library using primers AP1 and AP2. The forward primer AP1 was designed to anneal to nucleotide sequences of the partial cDNA (SEQ ID NO: 9) encoding part of the first zinc finger motif present in the DNA-binding domain. The reverse primer, AP2, was adapted from degenerate primers designed to anneal to nucleotide sequences complementary to those encoding an EcR ligand binding domain (Kamimura *et al.,* 1996). The nucleotide sequences of primers AP1 and AP2 are as follows:

Primer AP1: 5′- TCGTCCGGTTACCATTACAACGC -3′ (SEQ ID NO: 27); and
Primer AP2: 5′- TAGACCTTTGGC(A/G)AA(C/T)TC(A/G/C/T)ACAAT -3′(SEQ ID NO: 28)

The PCR reaction mixture contained 4 µl of each primer (50 pm/µl), 5 µl of deoxynucleotide triphosphate mix (2mM), 1 µl of aphid oligo dT primed Lambda ZAPII cDNA library, 1 µl of recombinant *Pfu* DNA Polymerase (5 units/µl, Stratagene®), 5 µl of 10x *Pfu* buffer (Stratagene®) and 30 µl of MilliQ water. The *Pfu* polymerase was used in this reaction because it possesses proof-reading activity, which reduces the possibility of misincorporation of nucleotides. The PCR conditions included 42 cycles, each cycle comprising annealing at 55°C, extension at 72°C and melting at 94°C.

The major amplification product obtained in this reaction was gel-purified, kinased and ligated into the *Sma*I site of pUC18.

To screen *M. persicae* cDNA libraries, the cloned amplification product was digested to generate two non-overlapping probes, designated "EcR probe 1" (i.e. SEQ ID NO: 11) and "EcR probe 2" (i.e. SEQ ID NO: 12). In this regard, digestion of the cloned product with *Sph*I produced a DNA fragment comprising a nucleotide sequence specific for a region encoding the

DNA-binding domain (EcR probe 1; SEQ ID NO: 11), whilst digestion with *SphI/EcoRI* produced a DNA fragment comprising a nucleotide sequence having homology to a region encoding a putative linker domain, designated domain D, and the 5′-end of a putative hormone-binding domain, present in the EcR polypeptide of the insect ecdysone receptors (EcR probe 2, SEQ ID NO: 12).

EcR probe 1 and EcR probe 2 were labelled with [α-^{32}P]dATP in a reaction catalysed by Klenow fragment. All reagents were components of a GIGAprime DNA labelling kit (BresaGen Limited, Adelaide, Australia), except that the random primers were replaced with specific oligonucleotides synthetisezed to be complementary to the ends of EcR probe 1 and EcR probe 2.

Screening of *M. persicae* cDNA libraries

480,000 plaques from the oligo dT primed Lambda Zap Express cDNA library (Example 7) were screened as described above, using EcR probe 1. This approach yielded about 300 positive clones. Positive-hybridising clones were pooled and rescreened separately using EcR probe 1 and EcR probe 2, on duplicate lifts. Only four plaques were identified which hybridised to both probes. One of these was found by sequencing to contain a full-length cDNA encoding the EcR polypeptide of the *M. persicae* ecdysone receptor. The nucleotide sequence of the open reading frame of this cDNA is set forth herein as SEQ ID NO: 9. The derived amino acid sequence of the EcR polypeptide subunit of the *M. persicae* ecdysone receptor encoded by this open reading frame is set out in SEQ ID NO: 10.

EXAMPLE 9

***In vivo* function of recombinant EcR polypeptides of the *L. cuprina* ecdysone receptor**

Construction of plasmid pF3

Plasmid pF3 was constructed in four steps as follows:

First, plasmid p5S1, comprising the full-length cDNA encoding the EcR polypeptide of the *L. cuprina* ecdysone receptor , was digested with *EarI* and a 3′ *EarI* cDNA fragment thus

generated, encoding the C-terminal end of the EcR polypeptide of the *L. cuprina* ecdysone receptor, was end-filled and sub-cloned into the HindII site of pUC19, to construct plasmid pEAR. In plasmid pEAR, the 3' end of the cDNA was oriented towards the KpnI site of the pUC19 vector.

Second, plasmid p5S1 was also digested separately with:

(1) *Apo*I and *Pst*I, to isolate the 5' end of the cDNA as a 179 bp fragment (fragment A);

(2) *Pst*I and *Spe*I, to isolate a 1650 bp cDNA fragment (fragment B); and

(3) *Spe*I and *Bgl*II, to isolate a 203 bp fragment (fragment C).

Third, plasmid pEAR was digested with *Bgl*II and *Kpn*I, to isolate the 3' end of the cloned cDNA fragment therein as a 313 bp fragment (fragment D).

Fourth, DNA fragments A, B, C and D were each isolated by agarose electrophoresis and ligated together into pBluescriptSK+, which had been digested with *Eco*RI and *Kpn*I, to produce plasmid pF3.

Plasmid pF3 thus contains the complete open reading frame of the cDNA encoding the EcR polypeptide of the *L. cuprina* ecdysone receptor, as a 2368 bp fragment located between two *Bam*HI sites.

<u>Construction of plasmid pSGLcEcR and plasmid pLcK8</u>

Plasmid pSGLcEcR was constructed by cloning the 2368 bp *Bam*HI fragment from pF3, into the *Bam*HI site of the mammalian expression vector pSG5 (Stratagene). Plasmid pLcK8 is a clone of pSGLcEcR.

<u>Construction of plasmid pSGDmEcR</u>

Plasmid pSGDmEcR is identical to plasmid pSV40-EcR (Example 1) comprising the EcR polypeptide of the *D. melanogaster* ecdysone receptor placed operably under control of the SV40 promoter.

Transfection of CHO cells

CHO cells were co-transfected with a mixture comprising the following DNAs, lysed and assayed for CAT and β-galactosidase enzyme activity, as described in the preceding Examples:

(1) one of the expression plasmids designated pSGDmEcR, or pSGLcEcR, or the parental expression plasmid pSG5 as a negative control, at a concentration of 1 μg/ml; and

(2) the CAT reporter plasmid $p(EcRE)_7$-CAT at a concentration of 1 μg/ml; and

(3) an independent LacZ reporter plasmid, pPGKLacZ, at a concentration of 1ug/ml, included as a control to monitor transfection efficiency .

CAT reporter gene expression was induced with 10 μM or 50 μM Muristerone A. In control samples, cells received only the carrier ethanol in place of Muristerone A.

ELISA was used to quantify the synthesis of CAT and β-galactosidase enzymes, in extracts of cells forty eight hours after transfection. Account was taken of the variation between experiments, by normalizing the level of CAT enzyme to the level of β-galactosidase enzyme present in the same extract. Fold induction represents the normalized values for CAT gene expression in cells transfected with pSGDmEcR, pSGLcEcR or pSG5 in the presence of hormone divided by the normalized values for CAT gene expression in cells transfected with the same plasmid but in the absence of hormone. The average values of three independent experiments are shown in Figure 1 and the error bars indicate standard error of the mean.

Data shown in Figure 1 indicate that the EcR polypeptide of the *L. cuprina* ecdysone receptor from Example 3 is biologically active *in vivo.* CAT induction is observed at both 50 μM and 10 μm steroid (Muristerone A), with about 30 and 15 fold induction respectively. In view of the *in vivo* activity of the EcR polypeptide of the *L. cuprina* ecdysone receptor obtained according to this protocol, potential insecticidal substances acting by interaction with an insect steroid receptor, such as an ecdysone receptor, are screened by addition of the substances to the *in vivo* assay described herein. Substances are added in an amount from 0.05 μM to 100 μM. Candidate insecticidal compounds are identified by their ability to modulate the reporter gene expression which results from trans-activation by the EcR polypeptide of the *L. cuprina* ecdysone receptor.

EXAMPLE 10

Chimeric EcR polypeptides of insect ecdysone receptors

Chimeric ecdysone receptors comprsing regions derived from EcR polypeptides of ecdysone receptors of different species are produced and assayed for enhanced activity. In a particularly preferred embodiment, a chimeric ecdysone receptor is produced using the EcR polypeptides of the *D. melanogaster, M. persicae* and *L. cuprina* ecdysone receptors.

In one exemplification of this embodiment, plasmids pSGLD and pSGDL are produced comprising coding regions derived from the EcR polypeptides of the *D. melanogaster* and *L. cuprina* ecdysone receptors. In plasmid pSGLD, the 5′-end of the open reading frame of the *D. melanogaster* sequence, encoding the N-terminal portion of the EcR polypeptide of the *D. melanogaster* ecdysone receptor to the end of the DNA-binding domain of said polypeptide, is fused to the 3′-end of the open reading frame of the *L. cuprina* sequence, encoding the C-terminal portion of the EcR polypeptide of the *L. cuprina* ecdysone receptor, from the D domain and hormone-binding domain to the carboxyl terminus. In plasmid pSGDL, the 5′-end of the open reading frame of the *L. cuprina* sequence, encoding the N-terminal portion of the EcR polypeptide of the *L. cuprina* ecdysone receptor to the end of the DNA-binding domain of said polypeptide, is fused to the 3′-end of the open reading frame of the *D. melanogaster* sequence, encoding the C-terminal portion of the EcR polypeptide of the *D. melanogaster* ecdysone receptor, from the D domain and hormone-binding domain to the carboxyl terminus. These plasmids thus encode chimeric EcR polypeptides which form ecdysone receptor variants.

As shown in Figure 2, chimeric EcR polypeptides of *L. cuprina* and *D. melanogaster* ecdysone receptors, comprising fusion polypeptides between the DNA-binding domains and hormone-binding domains of the base *L. cuprina* and *D. melanogaster* polypeptides, exhibit bioactivity when measured in the CAT assay described above. Significant bioactivity of the chimeric EcR polypeptides encoded by plasmids pSGLD and pSGDL, comparable to the bioactivity of the *D. melanogaster* base EcR polypeptide, is observed at both 10 μM and 50 μM concentrations of Muristerone A.

EXAMPLE 11

Isolation and characterisation of a full-length cDNA encoding the EcR partner protein (USP polypeptide) of the *L. cuprina* ecdysone receptor

The EcR partner protein (USP polypeptide) subunit of the *L. cuprina* ecdysone receptor also functions alone as a USP polypeptide of the *L. cuprina* juvenile hormone receptor. A cDNA encoding both receptor polypeptide activities was isolated using PCR and hybridisation as follows.

Hybridisation probe preparation

A 150 base-pair probe, specific for genetic sequences encoding the EcR partner protein (USP polypeptide) subunit of insect ecdysone receptors or the USP polypeptide subunit of insect juvenile hormone receptors (SEQ ID NO: 21), was isolated by PCR from *L. cuprina* genomic DNA using the degenerate primers described by Tzertzinis *et al.* (1994). The PCR reaction conditions were as described in Example 6, except that *Pfu* polymerase was used in place of *Taq*I polymerase.

The amplified DNA fragment was sub-cloned into *Eco*RI and *Cla*I double-digested *p*Bluescript SK+ vector (Stratagene), after double-digestion of the fragment using the enzymes *Eco*RI and *Cla*I, purification of the amplified fragment by agarose gel electrophoresis, and electro elution of the product band. The nucleotide sequence of the probe was obtained using the USB Sequenase version 2.0 Kit (SEQ ID NO: 21).

For probe preparation, the amplified *L. cuprina* DNA fragment was excised from the vector using *Eco*RI and *Sal*I, gel purified and ^{32}P-labelled using the GIGAprime DNA Labelling Kit (BresaGen Limited, Adelaide, Australia) essentially according to the manufacturer's instructions except that random primers were replaced with the two degenerate primers described by Tzertzinis *et al.* (1994) (see above). Unincorporated label was removed by size exclusion chromatography over Biogel-P60 (Biorad Ltd, Sydney, Australia). The probe was used at 10^6 cpm/ml in hybridizations.

Screening of *L. cuprina* cDNA libraries

The *L. cuprina* cDNA library described above (Example 6) was screened with the amplified probe as described in Example 6. From one positive plaque, we derived plasmid pLSP4 containing a 3800 bp insert. Sequencing revealed that the 5' portion of pLSP4 encodes the EcR partner protein (USP polypeptide) of the *L. cuprina* ecdysone receptor, followed by a long (2.4 kb), apparently untranslated region (UTR). A 2453 bp *Eco*RI fragment of plasmid pLSP4 was isolated and sub-cloned into pBluescript SK+ (Stratagene), to construct plasmid pBLU1, which contains the full-length cDNA sequence. The nucleotide sequence of the full-length cDNA present in pBLU1 and the amino acid sequence encoded therefor, are set forth herein as SEQ ID NO: 3 and SEQ ID NO: 4, respectively.

The open reading frame (ORF) of SEQ ID NO: 3 encodes a polypeptide comprising 467 amino acids in length. The ATG start codon is located within a very favourable translational start context (i.e. 5'-GAAA**ATG**-3') having 75% identity to the consensus sequence (i.e. 5'-C/AAAA**ATG** -3') for *D. melanogaster* mRNA sequences (Cavener *et al.*, 1987). Moreover, the derived amino acid sequence of the *L. cuprina* EcR partner protein (USP polypeptide) comprises domains A/B, C, D, and E/F that are characteristic of nuclear hormone receptors (Evans, 1988; Forman and Samuels, 1990).

The nucleotide sequences of the 5'- untranslated region and coding region of the cDNA contained in plasmid pLSP5, and the amino acid sequence encoded therefor, are set forth herein as SEQ ID NO: 5 and SEQ ID NO:.6, respectively. The nucleotide sequences of the 5'-untranslated region and coding region of the cDNA contained in plasmid pLSP12, and the amino acid sequence encoded therefor, are set forth herein as SEQ ID NO: 7 and SEQ ID NO: 8, respectively.

Nucleotide sequence analyses revealed differences in the 5'- untranslated regions of pLSP4, pLSP5, and pLSP12, however the coding regions appeared to be identical, suggesting a possible splice variation. This conclusion is supported by the fact that the cDNAs of pLSP4, pLSP5 and pLSP12 contained identical nucleotide sequences within their 5'- untranslated regions, however differed by the addition/deletion of sequences. In particular, the 5'-terminal 13 nucleotides of all three cDNA clones were identical, as was the nucleotide sequence

surrounding the translation start codon (i.e. 5'-AAA**ATG**-3'). Clone pLSP5 (SEQ ID NO: 5) differed from clone pLSP 4 (SEQ ID NO: 3) in so far as it included an additional 176 bp of 5'-untranslated sequence inserted between nucleotides 13 and 14 of pLSP4. Clone pLSP12 (SEQ ID NO: 7) also differed from pLSP4 (SEQ ID NO: 3) in so far as it included an additional 116 bp of 5'- untranslated sequence inserted between nucleotides 13 and 14 of pLSP4. Clones pLSP5 (SEQ ID NO: 5) and pLSP12 (SEQ ID NO: 7) differed in so far as pLSP5 included an additional 60 bp of 5'- untranslated sequence inserted between nucleotides 13 and 14 of pLSP12.

The ATG start codons of both clones pLSP5 and pLSP12 are within translational start context sequences (i.e. 5'-CAAA**ATG**-3') having absolute identity to the consensus sequence (i.e. 5'-C/AAAA**ATG** -3') for *D. melanogaster* mRNA sequences (Cavener *et al.*, 1987).

EXAMPLE 12

Isolation and characterisation of a partial cDNA encoding the EcR partner protein (USP polypeptide)of the *M. persicae* ecdysone receptor

The EcR partner protein (USP polypeptide) subunit of the *M. persicae* ecdysone receptor also functions alone as a USP polypeptide of the *M. persicae* juvenile hormone receptor. To isolate a partial cDNA encoding both receptor polypeptide activities, a 140 bp probe was amplified from *M. persicae* genomic DNA, by PCR, using the two degenerate primers described by Tzertzinis *et al.*(1994) (see preceding Example). The PCR reaction conditions were as described in Example 6, except that *Pfu* polymerase was used in place of *Taq*I polymerase.

The amplified DNA fragment was sub-cloned into *Eco*RI and *Cla*I double-digested *p*Bluescript SK+ vector (Stratagene), after double-digestion of the fragment using the enzymes *Eco*RI and *Cla*I, purification of the amplified fragment by agarose gel electrophoresis, and electro elution of the product band.

The nucleotide sequence of the insert in the *p*Bluescript SK+ vector was obtained using automated fluorescent dye terminator sequencing (SUPAMAC, Sydney Australia).

Hybridisation probe preparation and library screening

For probe preparation the amplified *M. persicae* DNA insert was cut out of the *p*Bluescript+ vector with *Eco*RI and *Sal*I, gel purified and ^{32}P-labelled using the GIGAprime DNA Labelling Kit (Bresatec Limited, Adelaide, Australia) essentially according to the manufacturer's instructions except that random primers were replaced with the degenerate primers described by Tzertzinis *et al.*(1994) (see preceding Example). Unincorporated label was removed by size exclusion chromatography over Biogel-P60 (Biorad Ltd, Sydney, Australia). The probe was used at 10^6 cpm/ml in hybridizations to screen the *M. persicae* cDNA library as described in Examples 7 and 8.

The positive-hybridising clones were plaque-purified and sequenced using standard procedures as described herein. The nucleotide sequence of the open reading frame of the full-length cDNA encoding the partner protein (USP polypeptide) subunit of the *M. persicae* ecdysone receptor or the USP polypeptide of the *M. persicae* juvenile hormone receptor is set forth herein as SEQ ID NO: 15. The derived amino acid sequence of this open reading frame is set forth as SEQ ID NO: 16.

EXAMPLE 13

A construct for the baculovirus-directed co-expression of functional ligand-binding regions of the EcR polypeptide and partner protein (USP polypeptide) of the *D. melanogaster* ecdysone receptor

A vector was prepared to facilitate the baculovirus-directed co- expression of ligand-binding regions derived from the EcR polypeptide and partner protein (USP polypeptide) of the *D. melanogaster* ecdysone receptor, the protein products of which associate on co-expression to form a functional hormone-binding complex. The associated proteins are then used in high through-put assays or three-dimensional structural analysis. We have found that the ligand-binding domain, together with most of the linker domain of the EcR polypeptide subunit and of the EcR partner protein (USP polypeptide), are sufficient to associate to form a functional hormone-binding complex.

1. Isolation of the ligand-binding region of the EcR polypeptide of the *D. melanogaster* ecdysone receptor.

A *Sac* I- *Hin*dIII fragment encoding most of the linker (domain D) and all of the ligand-binding domain (domains E and F) of the EcR polypeptide of the *Drosophila melanogaster* ecdysone receptor was excised from a plasmid comprising DNA encoding the complete EcR polypeptide (Koelle *et al.* 1991). The excised fragment was cloned into *Sac*I - *Hin*dIII-digested expression vector pQE31(Qiagen), to produce the plasmid vector pQE31DmECR.

2. Construction of a baculovirus expressing the ligand-binding regions of EcR and USP polypeptides

A baculovirus was constructed for the co-expression in insect cells of:

(i) a cDNA region comprising a nucleotide sequence which encodes at least the ligand-binding domain and much of the linker domain of the EcR polypeptide of the *D. melanogaster* ecdysone receptor isolated as described at paragraph (1) above; and

(ii) a cDNA region comprising a nucleotide sequence which encodes at least the ligand-binding domain and much of the linker domain of the partner protein (USP polypeptide) of the *D. melanogaster* ecdysone receptor.

To produce this baculovirus, a *Eco*R I - *Hin*dIII fragment was excised from pQE31DmECR, said fragment encoding an oligo-His tag, and most of the linker domain, together with all of the ligand-binding domain of EcR polypeptide. This *Eco*R I - *Hin*dIII fragment was ligated into *Eco*R I - *Hin*dIII cleaved pFastBacDUAL, to produce the plasmid pDmEcR.DUAL. To insert gene sequences specific for the partner protein (USP polypeptide), the *Hin*dII - *Nsi*I fragment encoding most of the linker and all of the ligand-binding domain of the partner protein (USP polypeptide) was excised from a full-length cDNA clone in plasmid pZ7-1 (supplied by Vince Henrich) and ligated into *Nco*I - *Nsi*I cleaved pDmEcR.DUAL. A nucleotide sequence encoding a "FLAG" peptide was subsequently incorporated upstream of, and in the same reading frame as, the nucleotide sequence encoding ligand-binding region of the partner protein (USP polypeptide), by ligation into the unique *Sma*I site, thereby producing the plasmid pDmEcR.USP.DUAL. Plasmids containing the FLAG-encoding nucleotide sequence in the correct orientation were selected by nucleotide sequence determination.

The segment of pDmEcR.USP.DUAL which encodes the tagged ligand-binding region of the EcR polypeptide and partner protein (USP polypeptide) sequences, placed operably under the control of polyhedrin and p10 promoters, respectively, was recombined into a baculovirus genome, by employing the Tn7 transposition system (Luckow *et al*, (1993). The polypeptide products were then co-expressed in insect Sf21 and Sf9 cells, where they associated into a functional complex.

Expression of the tagged ligand-binding regions of the EcR polypeptide and partner protein (USP polypeptide) sequences was examined by immunoblot analysis of extracts derived from insect Sf21 cells infected with the recombinant baculovirus, employing antibodies directed against the oligo-His and FLAG tags. This analysis detected bands on immunoblot analysis of approximately the predicted sizes for the expressed tagged ligand-binding regions of the EcR polypeptide and partner protein (USP polypeptide).

The protein detected by anti-oligo-His-antibodies was enriched by affinity purification on nickel-NTA resin (Qiagen), and the FLAG-labelled protein was affinity-purified using FLAG M2 Affinity Gel (Kodak). It was further demonstrated that the oligo-His-tagged EcR polypeptide and the FLAG-tagged EcR partner protein (USP polypeptide) bound as a hetero-oligomeric complex to FLAG M2 Affinity Gel (Kodak).

Furthermore, binding assays, performed using a modification of the method of Yund *et al* (1978), demonstrated a highly-significant increase in the binding of the a labelled ecdysone analogue, [^{3}H] ponasterone A, in cells infected by the recombinant baculovirus, compared to the binding observed for the naturally-occurring ecdysone holoreceptor in *L. cuprina* embryos. In contrast, cells infected by a control virus displayed neither antibody-positive bands on western analysis, nor specific binding of [^{3}H] ponasterone A, above background levels. These data indicate correct folding and association of the variant polypeptides comprising the ligand-binding regions of the *D. melanogaster* EcR polypeptide and *D. melanogaster* partner protein (USP polypeptide). The correctly-folded and associated complex formed by the truncated EcR polypeptide and truncated EcR partner protein (USP polypeptide), is used for X-ray and NMR structural analysis and for high-throughput screens.

EXAMPLE 14

Construct for the baculovirus-directed co-expression of functional ligand-binding regions of the EcR polypeptide and partner protein (USP polypeptide) of the *L. cuprina* ecdysone receptor

A vector for the baculovirus-directed co- expression of ligand-binding domains derived from the EcR polypeptide and partner protein (USP polypeptide) of the *L. cuprina* ecdysone receptor was prepared essentially as described in the preceding Example.

<u>1. Isolation of the ligand-binding region of the EcR polypeptide of the *L. cuprina* ecdysone receptor.</u>

A *Sph*I – *Kpn*I fragment encoding most of the linker (domain D) and all of the ligand-binding domain (domains E and F) of the EcR polypeptide of the *L. cuprina* ecdysone receptor was excised from a cDNA clone encoding the complete EcR polypeptide and cloned into the *Sph*I – *Kpn*I cleaved expression vector pQE32 (Qiagen), to produce the plasmid pQE32LcEcR.

<u>2. Isolation of the ligand-binding region of the partner protein (USP polypeptide) of the *L. cuprina* ecdysone receptor.</u>

A DNA fragment encoding most of the linker domain and all of the ligand-binding domain of the partner protein (USP polypeptide) of the *L. cuprina* ecdysone receptor was sub-cloned to produce the plasmid pBLU1.

<u>3. Construction of a baculovirus expressing the ligand-binding regions of *L. cuprina* EcR and USP polypeptides</u>

A baculovirus was constructed for the co-expression in insect cells of:

(i) a cDNA region comprising a nucleotide sequence which encodes at least the ligand-binding domain and much of the linker domain of the EcR polypeptide of the *L. cuprina* ecdysone receptor isolated as described at paragraph (1) above; and

(ii) a cDNA region comprising a nucleotide sequence which encodes at least the

ligand-binding domain and much of the linker domain of the partner protein (USP polypeptide) of the *L. cuprina* ecdysone receptor isolated as described at paragraph (2) above.

To produce this baculovirus, a *Eco*R I – *Pst*I fragment derived from plasmid pQE32LcEcR, encoding an oligo-His tag and most of the linker domain together with all of the ligand-binding domain of the *L. cuprina* EcR polypeptide was ligated into *Eco*RI– *Pst*I cleaved pFastBac.DUAL, to produce the plasmid pLcEcR.DUAL. An *Ava*II–*Eco*RV fragment, encoding most of the linker and all of the ligand-binding domain of *L. cuprina* partner protein (USP polypeptide) was excised from plasmid pBLU1 and ligated, together with a "FLAG" encoding sequence into the *Pvu*II site of pLcEcR.DUAL, to produce plasmid pLcEcR.USP.DUAL .

The segment of pLcEcR.USP.DUAL which encodes the tagged ligand-binding regions of the EcR polypeptide and partner protein (USP polypeptide) sequences, placed operably under the control of polyhedrin and p10 promoters, respectively, was recombined into a baculovirus genome, by employing the Tn7 transposition system (Luckow *et al*, (1993). The polypeptide products were then co-expressed in insect Sf21 and Sf9 cells, where they associated into a functional complex.

Expression was examined by immunoblot analysis. Antibodies directed against oligo-His and FLAG tags detected bands on immunoblot analysis of approximately the predicted sizes for the expressed EcR and USP polypeptide regions respectively, in extracts from insect Sf21 cells infected with the recombinant baculovirus. The protein detected by anti-oligo-His was greatly enriched utilising a nickel-NTA resin (Qiagen) and the FLAG-labelled protein purified on FLAG M2 Affinity Gel (Kodak). It was also demonstrated by immunoblot analysis that oligo-His-tagged *L. cuprina* truncated EcR polypeptides and FLAG-tagged *L. cuprina* truncated EcR partner protein (USP polypeptide) bind as a hetero-oligomeric complex to FLAG M2 Affinity Gel (Kodak).

Furthermore, binding assays, carried out by a modification of the method of Yund *et al* (1978), demonstrated a highly-significant increase in the binding of the tritiated ecdysone analogue,

ponasterone A, in cells infected by recombinant virus indicating correct folding and association of the two protein subunits (Figure 3), greater than that of the ecdysone holoreceptor in *L. cuprina* embryos. Cells infected by a control virus displayed neither antibody-positive bands on western analysis nor specific binding of tritiated hormone above background.

Expression of the tagged ligand-binding regions of the *L. cuprina* EcR polypeptide and partner protein (USP polypeptide) sequences was examined by immunoblot analysis of extracts derived from insect Sf21 cells infected with the recombinant baculovirus, employing antibodies directed against the oligo-His and FLAG tags. This analysis detected bands on immunoblot analysis of approximately the predicted sizes for the expressed tagged ligand-binding regions of the *L. cuprina* EcR polypeptide and partner protein (USP polypeptide).

The protein detected by anti-oligo-His-antibodies was enriched by affinity purification on nickel-NTA resin (Qiagen), and the FLAG-labelled protein was affinity-purified using FLAG M2 Affinity Gel (Kodak).

Furthermore, binding assays, performed using a modification of the method of Yund *et al* (1978), demonstrated a significant increase in the binding of the labelled ecdysone analogue, [^{3}H] ponasterone A, in cells infected by the recombinant baculovirus, compared to the binding observed for the naturally-occurring ecdysone holoreceptor in *L. cuprina* embryos (Figure 3). In contrast, cells infected by a control virus displayed neither antibody-positive bands on western analysis, nor specific binding of [^{3}H] ponasterone A, above background levels.

These data indicate correct folding and association of the variant polypeptides comprising the ligand-binding regions of the *L. cuprina* EcR polypeptide and *L. cuprina* partner protein (USP polypeptide). The correctly-folded and associated complex formed by the truncated EcR polypeptide and trucated EcR partner protein (USP polypeptide), is used for X-ray and NMR structural analysis and for high-throughput screens.

EXAMPLE 15

A construct for the expression of the ligand-binding region of the USP polypeptide of the *L. cuprina* juvenile hormone receptor

The donor plasmid pLcEcR.USP.DUAL (Example 14) was digested with *Bss*HII and *Pst*I to remove the *L. cuprina* EcR polypeptide-encoding segment therein, thereby leaving the tagged ligand-binding region of the *L. cuprina* USP polypeptide-encoding nucleotide sequence. The digested plasmid was blunt-ended using T4 DNA polymerase and Klenow polymerase, isolated by gel purification, and finally re-ligated to produce the plasmid pLcUSP.SINGLE.

To produce recombinant baculovirus capable of expressing the tagged ligand-binding regions of the USP polypeptide, the segment of pLcUSP.SINGLE encoding this polypeptide and the p10 promoter sequence to which said segment is operably connected, was recombined into a baculovirus genome employing the Tn7 transposition system (Luckow *et al.*, 1993). The polypeptide product is then expressed to form a functional juvenile hormone-binding polypeptide and preferably, a modulator of a juvenile hormone receptor. The correctly-folded truncated USP polypeptide is used for X-ray and NMR structural analysis and for high-throughput screens.

EXAMPLE 16

In-vitro Screening for the Detection of Insecticidal Compounds

The EcR partner protein (USP polypeptide) of the insect ecdysone receptor and USP polypeptide of the insect juvenile hormone receptor of the present invention, optionally associated with the EcR polypeptides of insect ecdysone receptors of the present invention as described in the preceding Examples, are coupled to pins according to the procedure of Geysen *et al.* (1987), and reacted with candidate insecticidal compounds, generally at a concentration in the range from about 0.05 µM to about 100 µM of the candidate compound. The binding of compounds is detected using standard procedures, and compounds having insecticidal activity are identified. Preferably, such compounds exhibit insecticidal activity against a range of insects, including diptera, hemiptera, coleoptera, ants, and moths, amongst

others. More preferably, the compounds will exhibit insecticidal activity against *L. cuprina*, *M. persicae, D. melanogaster,* scale insect, white fly, and leaf hopper, amongst others. In a particularly preferred embodiment, insecticidal compounds are specific to *L. cuprina* or *M. persicae* and close relatives thereof.

EXAMPLE 17

Cloned Myzus persicae EcR/USP complex binds ponasterone A *in vitro*.

In vitro-translated *Myzus persicae* EcR (MpEcR) polypeptide and an *in vitro*-translated *M.persicae* USP (MpUSP) polypeptide were produced labelled with [^{35}S]Methionine, using the Promega TNT-Coupled Reticulocyte Lysate System. Each batch of lysate contained 100-200 mg/ml of endogenous proteins (using BSA as a standard). The products were analysed by SDS-PAGE and radioautography. The results confirmed that the cloned cDNAs encode proteins of the sizes predicted from the length of putative open reading frames of the cDNAs present in plasmids pMpEcR and pMpUSP. The yields of EcR and USP were similar as assessed by SDS-PAGE.

In functional assays, DNA plasmids pMpEcR (AGAL Accession No. NM99/04567; 1 mg) or pMpUSP (AGAL Accession No. NM99/04568; 1 mg) or pMpUSP2 (AGAL Accession No. NM00/12581; 1 mg), which have been constructed using the vector pBK-CMV, and 1 ml of appropriate TNT RNA Polymerase were added to 48 ml of reaction mix which contained TNT Lysate, TNT Reaction Buffer, amino acid mixture, Rnasin Ribonuclease Inhibitor and nuclease-free water in volumes specified in the manufacturer's protocol. In control reactions, a Luciferase T3 control DNA (Promega) was used in place of pMpEcR or pMpUSP. T7 RNA Polymerase was used for transcription of the *M. persicae* EcR RNA from plasmid pMpEcR, whilst T3 RNA Polymerase was used for transcription of *M. persicae* USP RNA from the plasmid pMpUSP and the Luciferase T3 control DNA. The reactions were carried out for 90 minutes at 30°C .

The control reaction produces 150-500 ng of luciferase per 50 ml reaction.

The ecdysteroid binding activities of an *in vitro*-translated *Myzus persicae* EcR (MpEcR)

polypeptide and an *in vitro*-translated complex of the *M. persicae* EcR and USP polypeptides were produced from the RNAs using the TNT-Coupled Reticulocyte Lysate System (Promega). The mixtures were stored at -20°C overnight.

After thawing the translation products, 15 ml aliquots of the reaction mixture containing *M. persicae* EcR and USP polypeptides were combined to promote formation of the EcR/USP complex. For assays of individual proteins, 15 ml of the reaction mixture containing *M. persicae* EcR polypeptide or 15 ml of the reaction mixture containing *M. persicae* USP polypeptide was combined with 15 ml of control luciferase protein reaction mixture. Samples were each diluted to 435 ml with EcR40 buffer [40 mM KCl, 25 mM HEPES pH 7.0, 1 mM EDTA, 1mM DTT, BSA (0.5 mg/ml), 10% glycerol] to allow for triplicates in the ligand binding assay. A control reaction (Blank) was established which contained EcR40 buffer only. An aliquot (140 ml) of each diluted sample was incubated with tritiated ponasterone A (DuPont NEN, Batch Number 3281108) at a final concentration of 2.2 nM for 90 min at room temperature. After incubation, the ligand binding reactions were placed on ice. The samples were pipetted onto Whatman GF/C filters and incubated for 30 sec. The filters were then placed on a vacuum sinter, washed with 10 ml EcR40 buffer and transferred to scintillation vials. After adding 7 ml of InstaGel Plus to each vial, the contents were vortexed and left at room temperature until the filters became transparent. The receptor bound ligand was quantified using a TriCarb 2100TR scintillation counter.

The results depicted in Figure 4 indicate that significantly higher amounts of ponasterone A bind to the complex than to either the USP or EcR polypeptides alone.

EXAMPLE 18

In vivo function of a chimeric *L. cuprina* ecdysone receptor and a *L. cuprina* EcR partner protein (USP polypeptide)

<u>Construction of plasmid pSGLcUSP</u>

A 2453 bp fragment from the 5' end of clone pLSP4 (Example 11), containing nucleotide sequence encoding the *L. cuprina* EcR partner protein (USP polypeptide), was subcloned into

the *Eco*RI site of the mammalian expression vector pSG5 (Stratagene), to construct pSGLcUSP.

Construction of plasmid pVPLcEcR

Plasmid pVPLcEcR was constructed as follows:

To construct plasmid pMOD31, plasmid pNLVP16 (a gift from Dr G. Muscat) was digested with *Sal*I and *Xba*I, and re-ligated using a double-stranded oligonucleotide linker formed by annealing of the following complementary oligonucleotides:

SPX5: 5′-TCGACATATAACTTCGCTGCAGATGCATCCGAGCTCT-3′ (SEQ ID NO: 29); and
XPS3: 5′-CTAGAGCTCGGATGCATCTGCAGCGAAGTTATATG-3′ (SEQ ID NO: 30),

The A/B domain of pSGLcEcR was removed from the EcR-encoding cDNA by digestion using the restriction enzymes *Bam*HI and *Pst*I, and a 263 bp *Bgl*II/*Pst*I fragment of plasmid pMOD31, containing a VP16 activation domain (Triezenberg *et al*., 1988), was ligated in its place, to construct plasmid pVPLcEcR. Accordingly, plasmid pVPLcEcR contains nucleotide sequences encoding the ligand binding region of the *L. cuprina* EcR polypeptide placed operably in connection with the VP16 activation domain.

Transfection of CV1 Cells

CV1 cells were cotransfected with (i) plasmid pSGLcUSP or unmodified plasmid pSG5, at 1 μg/ml; (ii) plasmid pVPLcEcR or unmodified plasmid pSG5, at 0.2 μg/ml; (iii) the CAT reporter plasmid p(EcRE)$_7$-CAT (Example 9), at 1 μg/ml; and (iv) an independent LacZ reporter plasmid, pPGKLacZ (Example 9), at a concentration of 1ug/ml, included as a control to monitor transfection efficiency.

For induction experiments, the ecdysone analogue, 1mM ponasterone A was added to cells 6 hours after transfection. In control experiments, cells were treated only with carrier ethanol.

The CAT and β-galactosidase activities present in extracts of cells were measured 48 hours

after transfection as described previously (Hannan and Hill, 1997). Variations between experiments were controlled, by normalising the level of CAT to β-galactosidase for each extract.

Data shown in Figure 5 indicate that the *L. cuprina* EcR partner protein (USP polypeptide) can interact with a chimeric *L. cuprina* EcR polypeptide to form an ecdysteroid-dependent transcription factor in mammalian cells. Treatment of CV1 cells with ecdysteroid, in particular 1mM ponasterone A, induced significant levels of expression of the CAT reporter gene relative to the induction of β-galactosidase gene expression, which indicates transfection efficiency. When both plasmid pSGLcUSP and plasmid pVPLcEcR were transfected into CV1 cells, a 73-fold induction of CAT reporter gene expression, relative to β-galactosidase gene expression was achieved (column 8 of Figure 5).

In contrast, plasmid pVPLcEcR alone produced only a 4-fold induction of CAT gene expression relative to β-galactosidase gene expression (column 4 of Figure 5). This low level of activity is presumably due to formation of an active complex by the chimeric *L. cuprina* EcR polypeptide and endogenous RXR present in CV1 cells.

Only background CAT reporter gene expression was observed in the absence of exogenous hormone (columns 1, 3, 5 and 7 of Figure 5), and no significant induction of gene expression was observed in the absence of the *L. cuprina* EcR polypeptide (columns 1, 2, 5 and 6 of Figure 5).

Overall, these data support the conclusion that the cDNA clone described herein encodes an intact *L. cuprina* EcR partner protein (USP polypeptide), which is functional *in vivo*.

EXAMPLE 19

In vivo function of a chimeric EcR polypeptide of the *M. persicae* ecdysone receptor

Plasmids pSGDM and pSGMD both comprise nucleotide sequences encoding chimeric *D. melanogaster* and *M. persicae* ecdysone receptor EcR polypeptides. In particular, plasmid pSGDM comprises a chimeric cDNA sequence consisting of nucleotide sequence encoding the

A/B domain of the *D. melanogaster* EcR polypeptide ligated to nucleotide sequence encoding the DNA-binding, linker, and ligand-binding domains of the *M. persicae* EcR polypeptide. Plasmid pSGMD comprises a chimeric cDNA sequence consisting of nucleotide sequence encoding the A/B domain of the *M. persicae* EcR polypeptide ligated to nucleotide sequence encoding the DNA-binding, linker, and ligand-binding domains of the *D. melanogaster* EcR polypeptide.

Construction of plasmid pSGDM

cDNA encoding the *M. persicae* EcR polypeptide in the mammalian expression vector pSG5 (Stratagene) was digested with *Sap*I, which cleaves at a unique restriction site very close to the 3' -end of the nucleotide sequence encoding the A/B domain of the protein (Vector1). Two oligonucleotides A and B, containing *Sac*II, *Eco*RV and *Bam*HI restriction sites, were synthesized, purified and annealed to form double stranded DNA (Linker1) having *Sap*I compatible sticky ends:

A: 5'-TCCAGAACCGCGGATAGATATCTGGGATCCTC-3' (SEQ ID NO: 31); and

B: 5'-GGAGAGGATCCCAGATATCTATCCGCGGTTCT-3' (SEQ ID NO: 32)

Linker1 was ligated into Vector 1 and the resultant plasmid was digested with *Eco*RV to produce Vector 2.

A 940 bp *Eco*RI cDNA fragment encoding the A/B domain of full length *D. melanogaster* EcR polypypeptide was isolated from plasmid pSGDmEcR and ligated into the *Eco*RV site in Vector 2, using the linker-primer from the Stratagene cDNA Synthesis Kit, to produce Vector 3. The cDNA sequence encoding the A/B domain of the *M. persicae* EcR polypeptide was removed from Vector 3 by digestion with *Sac*II and the truncated plasmid was then religated to produce plasmid pSGDM.

Construction of plasmid pSGMD.

A 2200 bp *Eco*RI/*Bam*HI cDNA fragment encoding the DNA-binding and ligand-binding domains of the full length *D. melanogaster* EcR polypeptide was isolated from plasmid pSGDmEcR and end-filled and ligated into the *Eco*RV site present in Vector 2 (see above), to

produce Vector 4. The cDNA encoding the DNA-binding and ligand-binding domains of the *M. persicae* EcR polypeptide were then excised from Vector 4 by digestion with *Bam*HI and the truncated plasmid was then religated, to produce plasmid pSGMD.

Biological assays

Plasmids pSGDM and pSGMD contain cDNA sequences encoding full-length functional EcR polypeptides, as shown by SDS/PAGE of *in vitro* translation products, and using biological activity assays carried out *in vivo* using CHO cells, as follows:

1. Transfection of CHO cells.

CHO cells were co-transfected with a mixture comprising the following plasmids:

(i) an expression plasmid selected from the group consisting of pSGDmEcR, pSGMpEcR, pSGDM, pSGMD, and pSG5, wherein each plasmid was at a concentration of 1 µg/ml; and

(ii) the CAT reporter plasmid p(EcRE)$_7$-CAT at a concentration of 1 µg/ml.

Transfected cells were incubated for two days at 37°C in the presence of absence of 10 µM Muristerone A. In the control samples lacking Muristerone A, ethanol solvent was added to the cells. CAT enzyme activity was assayed by ELISA.

Data presented in Figure 6 show that the modified EcR subunit of the *M. persicae* ecdysone receptor is biologically active *in vivo*. The *M. persicae* EcR polypeptide having an A/B domain derived from *D. melanogaster* confers ecdysone responsiveness on CAT reporter gene expression in CHO cells, under the control of a promoter sequence containing the *D. melanogaster* hsp27 ecdysone response elements present in plasmid p(EcRE)$_7$-CAT.

EXAMPLE 20

Co-expression of the ligand binding region of the *M. persicae* EcR polypeptide and the ligand binding region of the *L. cuprina* EcR partner protein (USP polypeptide) produces an active heterodimeric complex

A vector for the baculovirus-directed co-expression of ligand-binding regions derived from the EcR protein and partner protein (USP polypeptide) of the Myzus persicae ecdysone receptor

was prepared in two stages:

First, cDNA encoding the linker region (domain D) and ligand-binding domain (domains E and F) of the *M. persicae* EcR polypeptide was cloned into the multiple cloning site of the plasmid pLcUSP.SINGLE (Example 15), in operable connection with the polyhedrin promoter sequence. Plasmid pLcUSP.SINGLE contains cDNA encoding linker and ligand binding domain of the *L. cuprina* partner protein (USP polypeptide) placed operably in connection with the p10 promoter. To achieve this end, pLcUSP.SINGLE was linearised using *Bam*HI and *Hin*dIII restriction enzymes, and ligated to a synthetic linker (i.e Linker 1) which was constructed by annealing the following oligonucleotides:

Oligonucleotide A:

5'-GATCCATGGGACACCATCACCATCACCATAGGCCTTCCGAACGCGGTGAATTCCGACA-3' (SEQ ID NO: 33);

Oligonucleotide B:

5'-AGCTTGTCGGAATTCACCGCGTTCGGAAGGCCTATGGTGATGGTGATGGTGTCCCATG-3' (SEQ ID NO: 34).

Linker 1 comprises *Bam*HI and *Hin*dIII sticky ends to facilitate cloning, and internal *Stu*I and *Eco*RI restriction sites, and nucleotide sequence encoding an oligo-His tag.

A 1.9kb *Stu*I/*Sma*I cDNA fragment encoding the linker and ligand-binding domain of the *M. persicae* EcR polypeptide was subsequently ligated into the *Stu*I restriction site within the Linker 1 sequence, to produce plasmid pMpEcR.LcUSP.DUAL, comprising nucleotide sequences encoding the tagged Linker 1 sequence, and domains D and E and F of the *M. persicae* EcR polypeptide, and the linker and ligand-binding domains (i.e. domains D/E/F) of the *L. cuprina* partner protein (USP polypeptide), placed operably under the control of polyhedrin and p10 promoters, respectively.

Second, plasmid pMpEcR.LcUSP.DUAL was digested with *Xma*I and *Kpn*I to excise nucleotide sequence encoding linker and ligand binding domains of the *L. cuprina* partner protein (USP polypeptide), to produce Vector 1B.

Linker 2 was constructed by annealing the following oligonucleotides:

Oligonucleotide C:

5'-CCGGGATCTCGAGATGGACTACAAGGACGACGATGACAAGCC-3' (SEQ ID NO: 35); and

Oligonucleotide D:

5'-CATGGGCTTGTCATCGTCGTCCTTGTAGTCCATCTCGAGATC-3' (SEQ ID NO: 36).

Linker 2 comprises *Xma*I and *Nco*I compatible ends and a "FLAG" encoding sequence.

Linker 2 was ligated to a 1.2kb *Kpn*I/*Nco*I DNA fragment in Vector 1B, encoding the linker and ligand binding domains of the *M. persicae* partner protein (USP polypeptide), to produce plasmid pMpEcR.USP.DUAL, comprising nucleotide sequences encoding the tagged linker and domains D and E and F of the *M. persicae* EcR polypeptide, and the linker and ligand-binding domains of the *M. persicae* partner protein (USP polypeptide), placed operably under the control of polyhedrin and p10 promoters, respectively.

Plasmids pMpEcR.LcUSP.DUAL and pMpEcR.USP.DUAL were sequenced to confirm the presence of the open reading frames.

The segment of pMpEcR.LcUSP.DUAL or pMpEcR.USP.DUAL encoding the chimeric tagged ligand binding regions of the receptor polypeptides was recombined in a baculovirus genome, by employing the Tn7 transposition system (Luckow *et al.*, 1993). The chimeric ligand binding regions of the recombinant ecdysone receptors were then expressed in insect Sf9 cells, where they associated into functional complexes.

Expression of the heterologous *M. persicae/L. cuprina* ecdysone receptor [i.e. comprising tagged linker and domains D/E/F of the *M. persicae* EcR polypeptide and the linker and ligand-binding domains of the *L. cuprina* EcR partner protein (USP polypeptide)], and expression of the homologous *M. persicae* ecdysone receptor [i.e. comprising tagged Linker 2 and domains D/E/F of the *M. persicae* EcR polypeptide and the linker and ligand-binding domains of the *M. persicae* EcR partner protein (USP polypeptide)], was examined by immunoblot analysis of extracts derived from insect Sf9 cells infected with either of the recombinant baculoviruses, employing antibodies directed against the oligo-His and FLAG tags to perform the quantitation. This analysis detected bands on immunoblot analysis of the predicted sizes for the expressed

polypeptides.

Furthermore, binding assays, carried out by a modification of the method of Yund *et al.*(1978), demonstrated a highly-significant binding of the tritiated ecdysone analogue, ponasterone A, in cells infected by recombinant viruses (Figure 7). Data presented in Figure 7 indicate correct folding and association of the components of both the heterologous and homologous truncated ecdysone receptors. These data indicate further that it is possible to produce functional heterologous chimeric receptors between the ligand binding regions of EcR polypeptides and EcR partner proteins from different insect species. Those chimeric receptors have different specificities for ecdysteroid compared to their native counterparts.

EXAMPLE 21

In vivo function of a heterodimeric receptor comprising a chimeric *M. persicae* EcR polypeptide and a recombinant *M. persicae* EcR partner protein (USP polypeptide

To test the function of the isolated cDNA clone encoding the *M. persicae* EcR partner protein (USP polypeptide), we tested the ability of the expressed polypeptide to complement a chimeric *M. persicae* EcR polypeptide in CV1 cells.

Briefly, CV1 cells were co transfected with the following plasmid constructs:

(i) plasmid pBKMpUSP1, containing the cDNA clone encoding the *M. persicae* EcR partner protein (USP polypeptide) operably in connection with the CMV promoter in pBK-CMV, at 2 µg/ml; or alternatively, a negative control gene construct, plasmid pBSK+, at 2 µg/ml; and

(ii) plasmid pSGDM, comprising a chimeric cDNA sequence consisting of nucleotide sequence encoding the a/B domain of the *D. melanogaster* EcR polypeptide ligated to nucleotide sequence encoding the DNA-binding, linker, and ligand-binding domains of the *M. persicae* EcR polypeptide (Example 19), at 1 µg/ml; and

(iii) the CAT reporter gene construct, plasmid p(EcRE)$_7$-CAT, comprising the CAT reporter gene placed operably under the control of a promoter sequence containing multiple copies of the *D. melanogaster* hsp27 ecdysone response elements present in plasmid at 1 µg/ml; and

(iv) the β-galactosidase reporter gene construct to control for transfection efficiency, designated plasmid pPopNLacZ, and described by Hannan *et al.* (1993), at 0.5 μg/ml.

For expression, the ecdysone analogue, 10 mM ponasterone a (a gift from Dr Denis Horn), was added to cells 6 hours after transfection, to induce CAT gene expression mediated the chimeric recombinant ecdysone receptor. In control experiments, cells were treated with ethanol in place of ponasterone a. CAT and β-galactosidase enzyme activities were measured in cell extracts 48 hours after transfection, as described previously (Hannan and Hill, 1997). The relative level of CAT/ β-galactosidase for each extract was determined, to normalise the variation in transfection efficiency between samples.

Data presented in Figure 8 indicate that the isolated cDNA encodes a functional *M. persicae* EcR partner protein (USP polypeptide). When pBKMpUSP1 was co transfected with pSGDM into CV1 cells, a 2.6-fold induction of relative CAT gene expression was observed in the presence of 10 mM ponasterone a, relative to the expression observed using plasmid pSGDM in the absence of pBKMpUSP1. The "background" level of gene expression observed for cells expressing plasmid pSGDM in the absence of pBKMpUSP1 is presumably due to formation of an active complex between the chimeric MpEcR polypeptide and the endogenous RXR proteins present in CV1 cells. The induction of CAT expression by ponasterone a for cells transfected with both plasmids pBKMpUSP1 and pSGDM indicates that the expressed *M. persicae* EcR partner protein (USP polypeptide) can interact with the chimeric EcR polypeptide, to form an ecdysteroid-dependent transcription factor in mammalian cells. Accordingly, these data indicate that the recombinantly-expressed *M. persicae* EcR partner protein (USP polypeptide) is functional *in vivo*.

EXAMPLE 22

Isolation and characterisation of a cDNA encoding the EcR partner protein (USP polypeptide) of the *Bemesia tabaci* ecdysone receptor

<u>Construction and screening of *B. tabaci* cDNA libraries</u>

Two independent *B. tabaci* cDNA libraries derived from red-eye nymph stage animals were prepared by oligo-dT priming, and cloned into the *Eco*RI site of the Lambda/ZapII vector

(Stratagene). The titres of the two primary libraries produced were 1.9 x 10^6 pfu, and 3.15 x 10^6 pfu. Tests indicated that the insert size range for these libraries was 0.7 kb to 7.6 kb in length.

The primary libraries generated were subsequently amplified according to the manufacturer's instructions, using standard protocols, to produce final titres of 1.5 x 10^9 pfu/ml, and 2.5 x 10^9 pfu/ml.

The prepared cDNA libraries were screened by lifting 500,000 plaques from each amplified cDNA library, in duplicate, onto Hybond N membranes (Amersham), and then hybridizing same under low stringency conditions to radiolabelled probes specified below. In particular, hybridisations were carried out overnight at 37°C, in a hybridisation solution comprising 42% (w/v) formamide; 5 x SSPE solution; 5 x Denhardt's solution; and 0.1% (w/v) sodium dodecyl sulphate, as described essentially by Ausubel *et al*, (1992) or Sambrook *et al.* (1989). The membranes were then washed at 37°C in 2XSSC solution containing 0.1% (w/v) sodium dodecyl sulphate. Following washing, positive plaques were detected by autoradiography, using XOMAT-AR film (Kodak) for two to three days, at -70°C. Positive-hybridising plaques were plaque-purified, rescued as plasmids, and their cDNA inserts analysed by nucleotide sequence determination.

Hybridisation probe preparation

The EcR partner protein (USP polypeptide) subunit of the *B. tabaci* ecdysone receptor also functions in the absence of the EcR polypeptide as a USP polypeptide of the *B. tabaci* juvenile hormone receptor.

To isolate a cDNA encoding both receptor activities from the *B. tabaci* cDNA library, a 140 bp probe was amplified from *B. tabaci* genomic DNA, using two degenerate primers described by Tzertzinis *et al.* (1994) and in the preceding Examples. The PCR reaction was performed using 1 unit *Taq*I polymerase (Boehringer Mannheim), 1 mM each primer, in a 50 µl reaction volume, essentially under conditions recommended by the manufacturer (Boehringer Mannheim).

The amplified DNA fragment was sub-cloned into the *Eco*RI and *Cla*I sites of linearised

pBluescript SK+ (Stratagene) vector. The nucleotide sequence of the insert in the pBluescript SK+ vector was obtained using automated fluorescent dye terminator sequencing (Automated DNA Analysis Facility at University of NSW, Sydney Australia) and is set forth herein as SEQ ID NO: 37. This fragment encodes the amino acid sequence set forth in SEQ ID NO: 38.

To prepare a hybridisation probe for screening cDNA libraries, the amplified *B. tabaci* DNA was released from the pBluescript+ vector by double-digestion using the enzymes *Eco*RI and *Sal*I, separation by agarose gel electrophoresis, and purification by electro-elution. DNA was subsequently [^{32}P]-labelled using the GIGAprime DNA Labelling Kit (Bresatec Limited, Adelaide, Australia) essentially according to the manufacturer's instructions, except that random primers were replaced with the degenerate primers described by Tzertzinis *et al.*(1994). Unincorporated label was removed by size exclusion chromatography over Biogel-P60 (Biorad Ltd, Sydney, Australia). The probe was used as described herein above at in Examples 7 and 8, to screen the *B. tabaci* cDNA library.

Positive-hybridising clones were plaque-purified and sequenced using standard procedures as described herein.

The nucleotide sequence of one clone was obtained and is set forth herein as SEQ ID NO: 39. The amino acid sequence of the *B. tabaci* EcR partner protein (USP polypeptide) is shown in SEQ ID NO: 40.

EXAMPLE 23

Cloning and characterization of a cDNA molecule encoding the EcR polypeptide of the *B. tabaci* ecdysone receptor

<u>Hybridisation probe preparation</u>

A 101 bp DNA fragment, encoding the DNA-binding domain of the EcR polypeptide subunit of the *B. tabaci* ecdysone receptor, was amplified from the *B. tabaci* genome by PCR, by using the degenerate primers Rdna3 (SEQ ID NO: 23) and Rdna4 (SEQ ID NO: 24), essentially as described hereinabove. Briefly, amplification reactions employed *Taq*I DNA polymerase (Boehringer Mannheim) and the following amplification conditions:

cycles 1-2: 97°C for 2 minutes; 50°C for 1 minute; 72°C for 1 minute;
cycles 3-30: 72°C for 7 minutes; and
cycle 44: 72°C for 7 minutes.

The amplified *B. tabaci* gene fragment was cloned into pBluescript SK+ (Stratagene), following digestion using the enzymes *Eco*RI and *Bam*HI, purification of the digested DNA by agarose gel electrophoresis and BresaClean (Bresatec)-purification of the product band, as described herein above.

The nucleotide sequence of the amplified hybridisation probe was obtained using automated fluorescent dye terminator sequencing (Automated DNA Analysis Facility at University of NSW, Sydney Australia), and is set forth herein as SEQ ID NO: 41. The derived amino acid sequence of this gene fragment is provided in SEQ ID NO: 42. There are 16 amino acids in the amino acid sequence of SEQ ID NO: 42 that are conserved in the amino acid sequence of the *M. persicae* EcR partner protein (USP polypeptide) set forth herein as SEQ ID NO: 16 (cf. SEQ ID NO: 42 to residues 63 to 95 of SEQ ID NO: 16), suggesting that the amplified probe does encode a part of the *B. tabaci* EcR partner protein (USP polypeptide).

For probe preparation, the insert was excised from the pBluescript SK+ vector using *Eco*R1 and *Bam*HI, labelled, and used as described in the preceding Example, to screen the *B. tabaci* cDNA library.

REFERENCES

1. Amann and Brosius (1985).*Gene* 40: 183.

2. Ausubel, *et al* (1992), *Current Protocols in Molecular Biology*, Greene/Wiley, New York

3. Bugg, *et al* (1993) *Scientific American*, December Issue: pp 60-66.

4. Cavener *et al.* (1987) *Nucl. Acids Res.* 15:1353-136.

5. Devereux, J., *et al.* (1984). *Nucl. Acids Res.* 12:387-395.

6. Evans (1988) *Science*, 240: 889-895.

7. Geysen *et al* (1987) *J. Immunol Methods*, 102, 259-274.

8. Greene *et al.* (1988) *Nucleic Acids Res.* 15:369.

9. Hannan et al (1993) Gene 130 : 233-239

10. Hollenberd and Evans (1988) *Cell* 55: 899-906.

11. Hollenberg *et al* (1991) *Cell*, 67, 59-77.

12. Jones, G and Sharp, P, (1997) *Proc. Nat. Acad. Sci. USA* 94:13499-13503.

13. Kamimura *et al* (1996) *Comp. Biochem. Physiol.* 113, 341-347.

14. Koelle *et al* (1991) *Cell*, 67:59-77

15. Krust *et al* (1986) *EMBO. J.*, 5: 891-897.

16. Luckow *et al*, (1993) *J. Virol*, 67: 4566.

17. McBurney *et al* (1991), *Nucleic Acids Res.*, 19, 5755-5761.

18. McPherson, M.J., Quirke, P. and Taylor, G.R. (1991) *PCR a Practical Approach.* IRL Press, Oxford University Press, Oxford, United Kingdom.

19. Needleman and Wunsch (1970) *J. Mol. Biol.* 48:443-453.

20. Potrykus (1990) *Bio/Technology* 8: 535-542.

21. Riddihough, G., and Pelham, H. R. B., (1987), *EMBO J.*, 6, 3729-3734

22. Sambrook *et al* (1989), *Molecular Cloning: a Laboratory Manual*, CSH Lab. Press

23. Segraves and Hogness (1990) *Genes Devel.* 4: 204-219.

24. Shimatake and Rosenberg (1981) *Nature* 292: 128

25. Studier and Moffat (1986) *J. Mol. Biol.* 189: 113

26. Thompson,*et al.,* (1994) *Nucl. Acids Res.* 22:4673-4680.

27. Tzertzinis *et al* (1994) *J. Mol. Biol.* 258, 479-486.

28. Von Itzstein *et al* (1993) *Nature* Vol. 363: 418-423.

29. Yund, *et al.* (1978) *Proc. Nat. Acad. Sci. USA*, 24: 6039-6043.

WE CLAIM:

1. An isolated nucleic acid molecule comprising a nucleotide sequence which encodes or is complementary to a sequence which encodes a steroid receptor polypeptide or a juvenile hormone receptor polypeptide or a bioactive derivative or analogue thereof, wherein said polypeptide is selected from the groups consisting of (i) an EcR polypeptide of a steroid receptor; (ii) the partner protein (USP polypeptide) of a steroid receptor; and (iii) the USP polypeptide of a juvenile hormone receptor; and wherein said polypeptide comprises an amino acid sequence having at least 40% identity to an amino acid sequence selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 38, SEQ ID NO: 40, and SEQ ID NO: 42 or having at least 40% identity to an amino acid sequence encoded by a cDNA present in a plasmid selected from the group consisting of AGAL Accession Nos. NM99/04565, NM99/04566, NM99/04567, NM99/04568, NM00/12580, and NM00/12581.

2. The isolated nucleic acid molecule according to claim 1, wherein the steroid receptor is an ecdysteroid receptor.

3. The isolated nucleic acid molecule according to claim 2, wherein the ecdysteroid receptor is an insect ecdysone receptor.

4. The isolated nucleic acid molecule according to claim 3, wherein the insect ecdysone receptor comprises the EcR polypeptide of an insect ecdysone receptor or the partner protein (USP polypeptide) of an insect ecdysone receptor.

5. The isolated nucleic acid molecule according to claim 4, wherein the insect is selected from the group consisting of diptera, hemiptera, coleoptera, lepidoptera, neuroptera, and ants.

6. The isolated nucleic acid molecule according to claim 5, wherein the hemipteran insect is *Myzus persicae* or a close relative thereof.

7. The isolated nucleic acid molecule according to claim 6, wherein the insect steroid

receptor polypeptide comprises an EcR polypeptide of the *M. persicae* ecdysone receptor having the amino acid sequence set forth in SEQ ID NO: 10 or SEQ ID NO: 14 or encoded by the cDNA present in plasmid pMpEcR (AGAL Accession No. NM99/04567) or a bioactive analogue or derivative thereof.

8. The isolated nucleic acid molecule according to claim 6, wherein the insect steroid receptor polypeptide comprises an EcR partner protein (USP polypeptide) of the *M. persicae* ecdysone receptor or a USP polypeptide of the *M. persicae* juvenile hormone receptor having or including an amino acid sequence selected from the group consisting of SEQ ID NO: 16, SEQ ID NO: 18, and SEQ ID NO: 20, or an amino acid sequence encoded by cDNA present in a plasmid selected from the group consisting of pMpUSP (AGAL Accession No. NM99/04568) and pMpUSP2 (AGAL Accession No. NM00/12581), or a bioactive analogue or derivative of any one of said amino acid sequences.

9. The isolated nucleic acid molecule according to claim 5, wherein the dipteran insect is *L. cuprina* or a close relative thereof.

10. The isolated nucleic acid molecule according to claim 9, wherein the insect steroid receptor polypeptide comprises an EcR polypeptide of the *L. cuprina* ecdysone receptor having the amino acid sequence set forth in SEQ ID NO: 2 or a sequence encoded by cDNA present in plasmid pLcEcR (AGAL Accession No. NM99/04566) or a bioactive analogue or derivative of said sequence.

11. The isolated nucleic acid molecule according to claim 9, wherein the insect steroid receptor polypeptide comprises an EcR partner protein (USP polypeptide) of the *L. cuprina* ecdysone receptor or a USP polypeptide of the *L. cuprina* juvenile hormone receptor having an amino acid sequence selected from the group consisting of SEQ ID NO: 4, SEQ ID NO: 6, and SEQ ID NO: 8, or an amino acid sequence encoded by cDNA present in plasmid pLcUSP (AGAL Accession No. NM99/04565) or a bioactive analogue or derivative of any one of said sequences.

12. The isolated nucleic acid molecule according to claim 5, wherein the insect is *Bemesia tabaci* (Silverleaf whitefly) or a close relative thereof.

13. The isolated nucleic acid molecule according to claim 12, wherein the insect steroid receptor polypeptide comprises an EcR partner protein (USP polypeptide) of the *B. tabaci* ecdysone receptor or a USP polypeptide of the *B. tabaci* juvenile hormone receptor having or including an amino acid sequence selected from the group consisting of SEQ ID NO: 37, and SEQ ID NO: 39, or an amino acid sequence encoded by cDNA present in plasmid pBtUSP (AGAL Accession No. NM00/12580), or a bioactive analogue or derivative of any one of said amino acid sequences.

14. The isolated nucleic acid molecule according to claim 12, wherein the insect steroid receptor polypeptide comprises an EcR polypeptide of the *B. tabaci* ecdysone receptor comprising the amino acid sequence set forth in SEQ ID NO: 42.

15. The isolated nucleic acid molecule according to claim 1, wherein the bioactive derivative or analogue comprises a fragment of an EcR polypeptide of an insect ecdysone receptor or a fragment of an EcR partner protein (USP polypeptide) of an insect ecdysone receptor, wherein said fragment includes at least one ligand-binding region of said EcR polypeptide or said EcR partner protein (USP polypeptide).

16. The isolated nucleic acid molecule according to claim 15, wherein the ligand-binding region comprises a linker domain of the EcR polypeptide or a linker domain of the EcR partner protein (USP polypeptide).

17. The isolated nucleic acid molecule according to claim 15, wherein the ligand-binding region comprises a hormone-binding domain of the EcR polypeptide or a hormone-binding domain of the EcR partner protein (USP polypeptide).

18. The isolated nucleic acid molecule according to claim 15, wherein the ligand-binding region comprises at least a part of the linker domain and all of the hormone-binding domain of

the EcR polypeptide or a linker domain and hormone-binding domain of the EcR partner protein (USP polypeptide).

19. The isolated nucleic acid molecule according to claim 1, comprising a protein-encoding nucleotide sequence having at least 40% identity to a nucleotide sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 37, SEQ ID NO: 39, and SEQ ID NO: 41, or having at least 40% identity to a complementary nucleotide sequence to any one of said sequences, or having at least 40% identity to a cDNA sequence present in a plasmid selected from the group consisting of AGAL Accession Nos. NM99/04565, NM99/04566, NM99/04567, NM99/04568, NM00/12580, and NM00/12581.

20. An isolated nucleic acid molecule comprising a nucleotide sequence which encodes or is complementary to a sequence which encodes a steroid receptor polypeptide or a juvenile hormone receptor polypeptide or a bioactive derivative or analogue thereof, wherein said nucleotide sequence is selected from the group consisting of:

(i) a nucleotide sequence having at least 40% identity to a nucleotide sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 37, SEQ ID NO: 39, and SEQ ID NO: 41, or a complementary nucleotide sequence to any one of said sequences;

(ii) a nucleotide sequence that is capable of hybridising under at least low stringency conditions to a nucleotide sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 37, SEQ ID NO: 39, and SEQ ID NO: 41, or a complementary nucleotide sequence to any one of said sequences;

(iii) a nucleotide sequence having at least 40% identity to a nucleotide sequence of a cDNA present in a plasmid selected from the group consisting of AGAL Accession Nos. NM99/04565, NM99/04566, NM99/04567, NM99/04568, NM00/12580, and

NM00/12581;

(iv) a nucleotide sequence that is capable of hybridising under at least low stringency conditions to cDNA present in a plasmid selected from the group consisting of AGAL Accession Nos. NM99/04565, NM99/04566, NM99/04567, NM99/04568, NM00/12580, and NM00/12581; and

(v) a nucleotide sequence that is amplifiable by PCR using a nucleic acid primer sequence selected from the group consisting of SEQ ID NO: 23, SEQ ID NO: 24, SEQ ID NO: 25, SEQ ID NO: 26, SEQ ID NO: 27, SEQ ID NO: 28, SEQ ID NO: 31, and SEQ ID NO: 32.

21. An isolated nucleic acid molecule which encodes an insect steroid receptor polypeptide and comprises the nucleotide sequence set forth in SEQ ID NO: 1 or a complementary nucleotide sequence thereto or the nucleotide sequence of the cDNA present in plasmid pLcEcR (AGAL Accession No. NM99/04566) .

22. An isolated nucleic acid molecule which encodes an insect steroid receptor polypeptide or a juvenile hormone receptor polypeptide and comprises a nucleotide sequence selected from the group consisting of SEQ ID NO: 3, SEQ ID NO: 5, and SEQ ID NO: 7, or a complementary nucleotide sequence to any one of said sequences, or the nucleotide sequence of a cDNA present in plasmid pLcUSP (AGAL Accession No. NM99/04565).

23. An isolated nucleic acid molecule which encodes an insect steroid receptor polypeptide and comprises a nucleotide sequence selected from the group consisting of SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 12, and SEQ ID NO: 13, or a complementary nucleotide sequence to any one of said sequences, or the nucleotide sequence of the cDNA present in plasmid pMpEcR (AGAL Accession No. NM99/04567).

24. An isolated nucleic acid molecule which encodes an insect steroid receptor polypeptide or a juvenile hormone receptor polypeptide and comprises a nucleotide sequence selected from the group consisting of SEQ ID NO: 15, SEQ ID NO: 17 and SEQ ID NO: 19, or a complementary nucleotide sequence to any one of said sequences, or the nucleotide sequence

of the cDNA present in plasmid pMpUSP (AGAL Accession No. NM99/04568) or plasmid pMpUSP2 (AGAL Accession No. NM00/12581).

25. An isolated nucleic acid molecule which encodes an insect steroid receptor polypeptide or a juvenile hormone receptor polypeptide and comprises the nucleotide sequence set forth in SEQ ID NO: 37 or SEQ ID NO: 39, or a complementary nucleotide sequence thereto, or the nucleotide sequence of the cDNA present in plasmid pBtUSP (AGAL Accession No. NM00/12580).

26. An isolated nucleic acid molecule which encodes an insect steroid receptor polypeptide and comprises a nucleotide sequence set forth in SEQ ID NO: 41.

27. a method of identifying an isolated nucleic acid molecule which encodes an insect steroid receptor polypeptide which encodes a steroid receptor polypeptide or a juvenile hormone receptor polypeptide comprising:

(i) hybridising genomic DNA, mRNA or cDNA with a hybridisation-effective amount of one or more probes selected from the group consisting of:

(a) a probe comprising at least 10 contiguous nucleotides in length from a nucleotide sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 24, SEQ ID NO: 25, SEQ ID NO: 26, SEQ ID NO: 27, SEQ ID NO: 28, SEQ ID NO: 31, SEQ ID NO: 32, SEQ ID NO: 37, SEQ ID NO: 39, and SEQ ID NO: 41, or a complementary nucleotide sequence to any one of said sequences;

(b) a probe comprising at least 10 contiguous nucleotides in length from cDNA contained in a plasmid selected from the group consisting of AGAL Accession Nos. NM99/04565, NM99/04566, NM99/04567, NM99/04568, NM00/12580, and NM00/12581; and

(c) a hybridisation probe comprising a nucleotide sequences selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID

NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 37, SEQ ID NO: 39, and SEQ ID NO: 41, or a complementary nucleotide sequence to any one of said sequences, or a homologue, analogue or derivative thereof having at least 40% identity to any one of said sequences or complementary sequences; and

(ii) detecting the hybridisation.

28. The method of claim 27 wherein the step of detecting the hybridisation comprises detecting a reporter molecule that is covalently bound to the probe.

29. a method of identifying an isolated nucleic acid molecule which encodes an insect steroid receptor polypeptide which encodes a steroid receptor polypeptide or a juvenile hormone receptor polypeptide comprising:

(i) annealing to genomic DNA, mRNA or cDNA, one or more PCR primers comprising at least 10 contiguous nucleotides in length derived from the group consisting of:

(a) a primer comprising at least 10 contiguous nucleotides in length from a nucleotide sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 24, SEQ ID NO: 25, SEQ ID NO: 26, SEQ ID NO: 27, SEQ ID NO: 28, SEQ ID NO: 31, SEQ ID NO: 32, SEQ ID NO: 37, SEQ ID NO: 39, and SEQ ID NO: 41, or a complementary nucleotide sequence to any one of said sequences;

(b) a primer comprising at least 10 contiguous nucleotides in length from cDNA contained in a plasmid selected from the group consisting of AGAL Accession Nos. NM99/04565, NM99/04566, NM99/04567, NM99/04568, NM00/12580, and NM00/12581; and

(ii) amplifying a nucleotide sequence which encodes a steroid receptor polypeptide or a juvenile hormone receptor polypeptide in a polymerase chain reaction.

30. a method of identifying an isolated nucleic acid molecule which encodes an insect steroid receptor polypeptide which encodes a steroid receptor polypeptide or a juvenile hormone receptor polypeptide comprising:

(i) amplifying a nucleotide sequence which encodes a steroid receptor polypeptide or a juvenile hormone receptor polypeptide in a polymerase chain reaction using one or more PCR primers selected from the group consisting of:

(a) a primer comprising at least 10 contiguous nucleotides in length from a nucleotide sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 24, SEQ ID NO: 25, SEQ ID NO: 26, SEQ ID NO: 27, SEQ ID NO: 28, SEQ ID NO: 31, SEQ ID NO: 32, SEQ ID NO: 37, SEQ ID NO: 39, and SEQ ID NO: 41, or a complementary nucleotide sequence to any one of said sequences; and

(b) a primer comprising at least 10 contiguous nucleotides in length from cDNA contained in a plasmid selected from the group consisting of AGAL Accession Nos. NM99/04565, NM99/04566, NM99/04567, NM99/04568, NM00/12580, and NM00/12581;

(ii) hybridising the amplified nucleotide sequence to genomic DNA, mRNA or cDNA with a hybridisation-effective amount of one or more probes selected from the group consisting of:

(a) a probe comprising at least 10 contiguous nucleotides in length derived from a sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 24, SEQ ID NO: 25, SEQ ID NO: 26, SEQ ID NO: 27, SEQ ID NO: 28, SEQ ID NO: 31, SEQ ID NO: 32, SEQ ID NO: 37, SEQ ID NO: 39, and SEQ ID NO: 41, or a complementary nucleotide sequence to any one of said sequences;

(b) a probe comprising at least 10 contiguous nucleotides in length derived

from a cDNA contained in a plasmid selected from the group consisting of AGAL Accession Nos. NM99/04565, NM99/04566, NM99/04567, NM99/04568, NM00/12580, and NM00/12581; and

(c) a hybridisation probe comprising a nucleotide sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 37, SEQ ID NO: 39, and SEQ ID NO: 41, or a complementary nucleotide sequence to any one of said sequences, or a homologue, analogue or derivative of any one of said sequences or complementary sequences having at least 40% identity to that sequence or complementary sequence; and

(iii) detecting the hybridisation.

31. The method of claim 30 wherein the step of detecting the hybridisation comprises detecting a reporter molecule that is covalently bound to the probe.

32. The method according to claim 30, further comprising the step of isolating the identified nucleic acid molecule.

33. A genetic construct comprising the isolated nucleic acid molecule according to claim 1 operably linked to a promoter sequence.

34. The genetic construct according to claim 33, wherein the promoter is selected from the group consisting of SV40, MMTV, polyhedron and p10 promoters.

35. A recombinant or isolated polypeptide comprising a steroid receptor polypeptide or juvenile hormone receptor polypeptide or a bioactive derivative or analogue thereof, wherein said polypeptide:

(i) is selected from the group consisting of the EcR polypeptide of a steroid receptor, the partner protein (USP polypeptide) of a steroid receptor and the USP polypeptide of a juvenile hormone receptor; and

(ii) comprises an amino acid sequence having at least 40% identity to an amino acid sequence selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 38, SEQ ID NO: 40, and SEQ ID NO: 42, or having at least 40% identity to an amino acid sequence encoded by a cDNA present in a plasmid selected from the group consisting of AGAL Accession Nos. NM99/04565, NM99/04566, NM99/04567, NM99/04568, NM00/12580, and NM00/12581;

wherein said polypeptide is substantially free of naturally-associated cellular components.

36. A cell comprising the nucleic acid molecule according to claim 1.

37. a cell comprising the genetic construct according to claim 33.

38. a cell which expresses the isolated or recombinant polypeptide according to claim 35.

39. a method of identifying a modulator of steroid receptor-mediated gene expression or juvenile hormone receptor-mediated gene expression comprising:

(i) assaying the expression of a reporter gene in the presence of the recombinant or isolated polypeptide according to claim 35 and a potential modulator; and

(ii) assaying the expression of the reporter gene in the presence of the recombinant or isolated polypeptide according to claim 35 and without said potential modulator; and

(ii) comparing expression of the reporter gene in the presence of the potential modulator to the expression of a reporter gene in the absence of the potential modulator,

wherein said reporter gene is placed operably under the control of a steroid response element (SRE) to which said steroid receptor binds or a promoter sequence comprising said SRE.

40. a method of identifying a potential insecticidal compound comprising:

(i) assaying the binding directly or indirectly of the recombinant or isolated polypeptide according to claim 35 to a steroid response element (SRE) to which said

polypeptide binds, in the presence of a candidate compound; and

(ii) assaying the binding directly or indirectly of the recombinant or isolated polypeptide according to claim 35 to a steroid response element (SRE) to which said polypeptide binds, in the absence of said candidate compound; and

(ii) comparing the binding assayed at (i) and (ii), wherein a difference in the level of binding indicates that the candidate compound possesses potential insecticidal activity.

41. a method of identifying a candidate insecticidally-active agent comprising the steps of:

a) expressing an EcR polypeptide of an insect steroid receptor or a fragment thereof which includes the ligand-binding region, optionally in association with an EcR partner protein (USP polypeptide) of an insect steroid receptor or ligand binding domain thereof, optionally in association with an insect steroid or analogue thereof so as to form a complex;

b) purifying or precipitating the complex;

c) determining the three-dimensional structure of the ligand binding domain of the complex; and

d) identifying compounds which bind to or associate with the three-dimensional structure of the ligand binding domain, wherein said compounds represent candidate insecticidally-active agents.

42. A method of identifying a candidate insecticidally-active agent comprising the steps of:

a) expressing a USP polypeptide of a juvenile hormone receptor or a fragment thereof which includes the ligand-binding region, optionally in association with an EcR polypeptide of an insect steroid receptor or ligand binding domain thereof, and optionally in association with an insect steroid or analogue thereof, so as to form a complex;

b) purifying or precipitating the complex;

c) determining the three-dimensional structure of the ligand binding domain of the complex; and

d) identifying compounds which bind to or associate with the three-dimensional

structure of the ligand binding domain, wherein said compounds represent candidate insecticidally-active agents.

43. A synthetic compound which interacts with the three dimensional structure of a polypeptide or protein selected from the group consisting of:

(i) an EcR polypeptide of a steroid receptor or a fragment or bioactive derivative thereof;

(ii) an EcR partner protein (USP polypeptide) of a steroid receptor or a fragment or bioactive derivative thereof;

(iii) a USP polypeptide of a juvenile hormone receptor or a fragment or bioactive derivative thereof; and

(iv) a functional receptor or protein complex formed by association of (i) and (ii),

wherein said compound is capable of binding to said polypeptide or protein to agonise or antagonise the binding activity or bioactivity thereof.

44. A method of identifying a synthetic compound for insecticidal activity comprising contacting the recombinant or isolated polypeptide according to claim 35 with said compound for a time and under conditions sufficient for binding to occur and detecting said binding using a detection means, wherein the occurrence of binding is indicative of potetnial insecticidal activity of the compound.

1/8

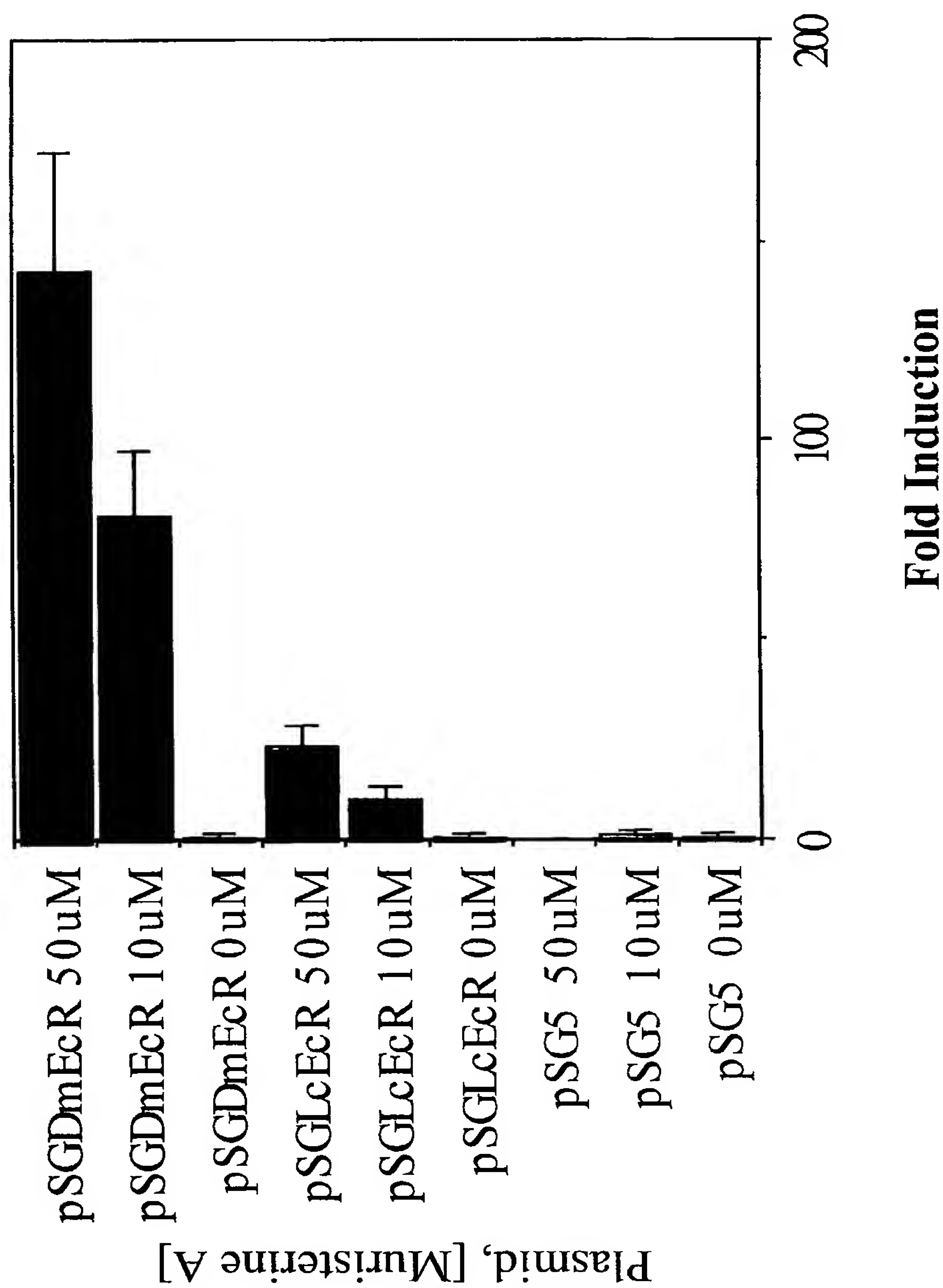

pSGDmEcR 50uM
pSGDmEcR 10uM
pSGDmEcR 0uM
pSGLcEcR 50uM
pSGLcEcR 10uM
pSGLcEcR 0uM
pSG5 50uM
pSG5 10uM
pSG5 0uM
Plasmid, [Muristerine A]
0
100
200
Fold Induction

FIGURE 1

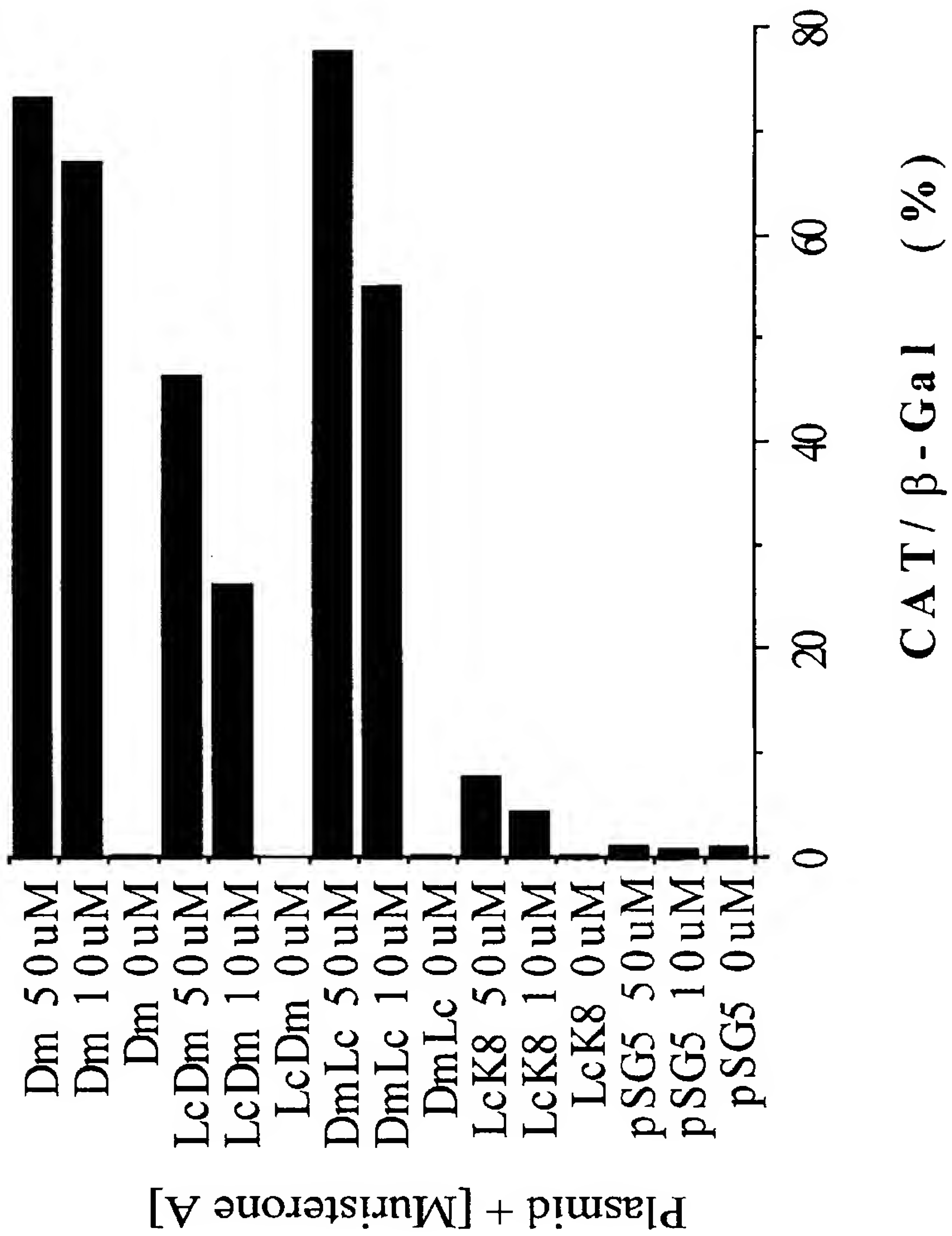

Plasmid + [Muristerone A]
Dm 50uM
Dm 10uM
Dm 0uM
LcDm 50uM
LcDm 10uM
LcDm 0uM
DmLc 50uM
DmLc 10uM
DmLc 0uM
LcK8 50uM
LcK8 10uM
LcK8 0uM
pSG5 50uM
pSG5 10uM
pSG5 0uM
0
20
40
60
80
CAT/β-Gal (%)

FIGURE 2

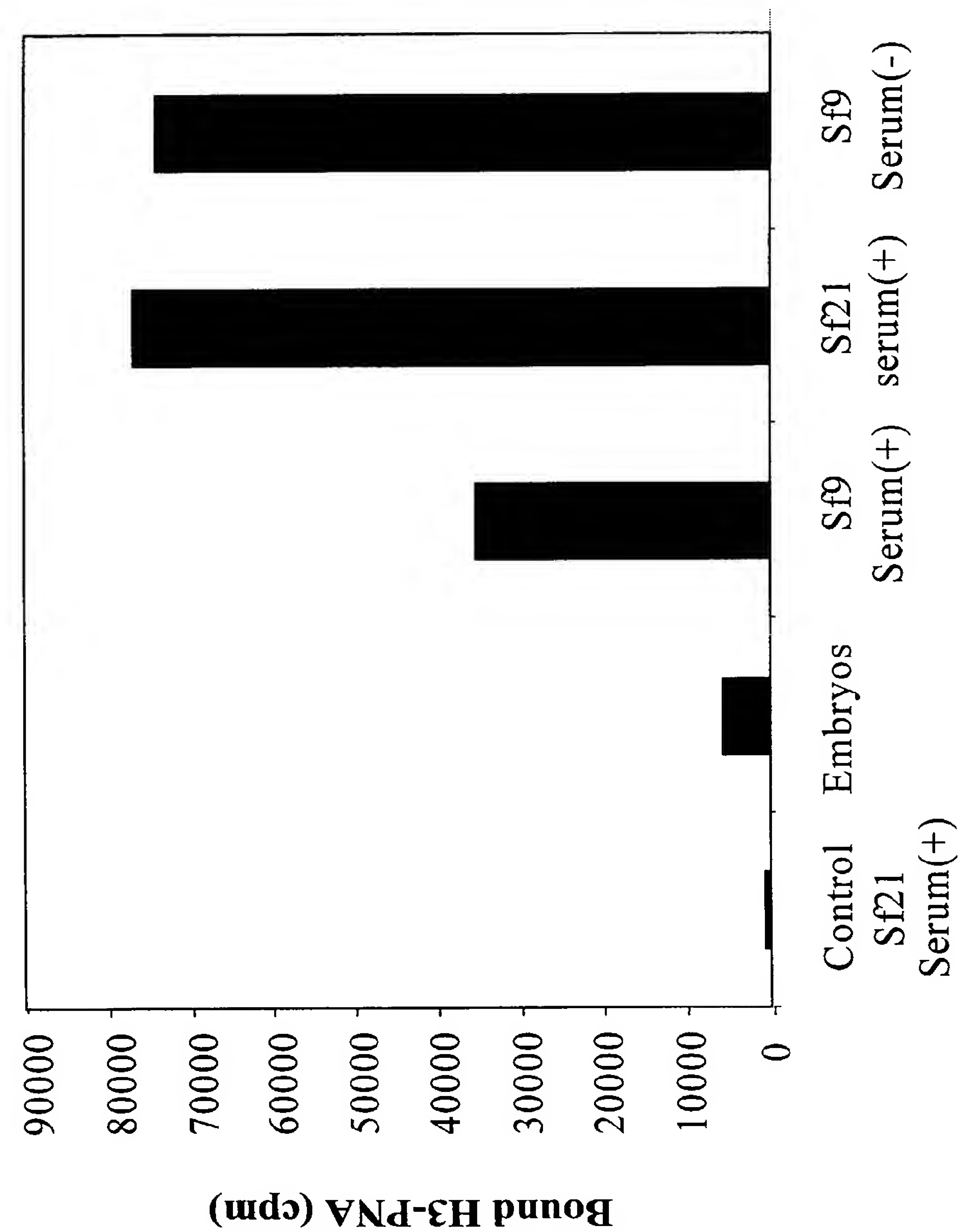

Bound H3-PNA (cpm)
90000
80000
70000
60000
50000
40000
30000
20000
10000
0
Control Sf21 Serum(+)
Embryos
Sf9 Serum(+)
Sf21 serum(+)
Sf9 Serum(-)

FIGURE 3

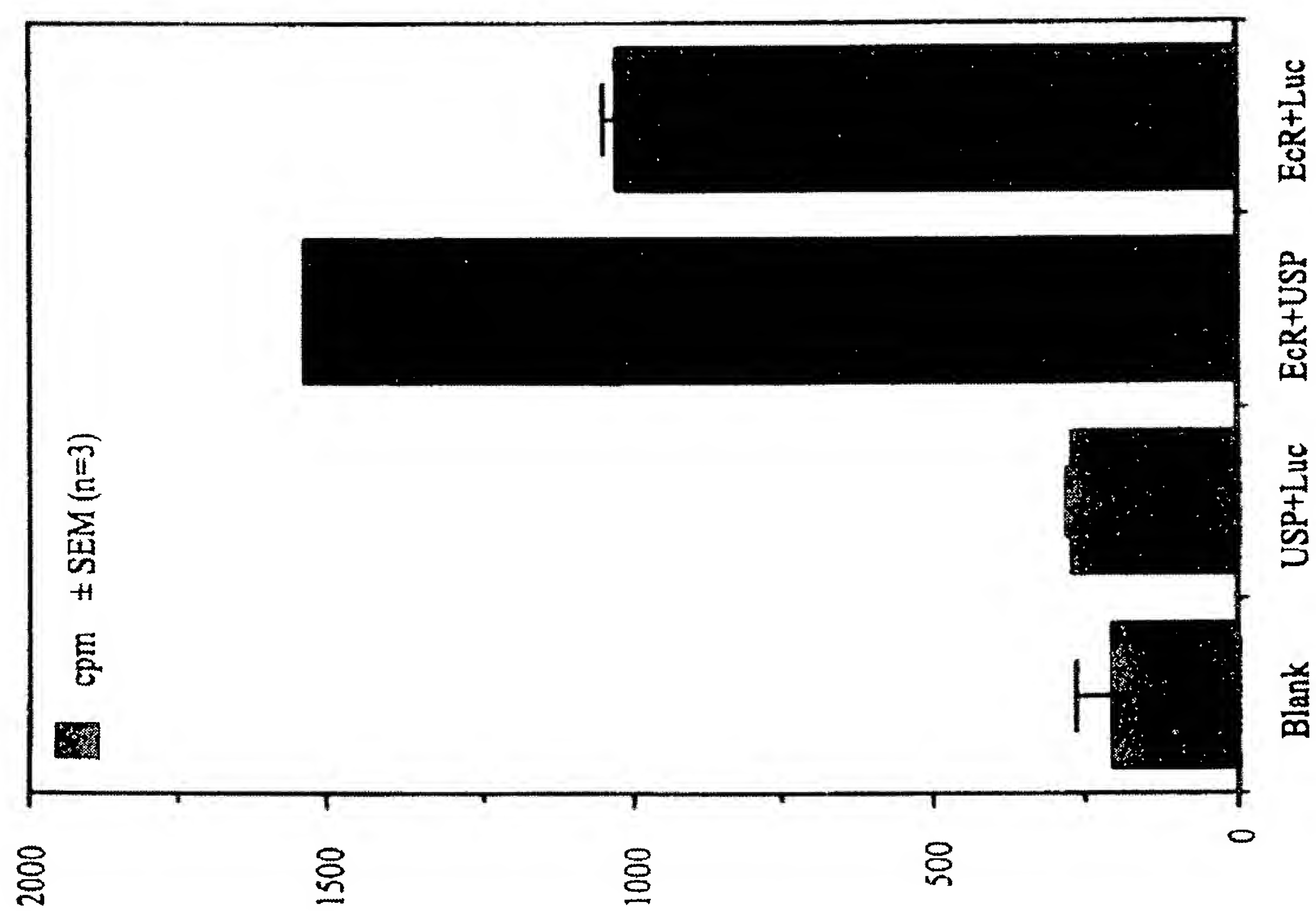


FIGURE 4

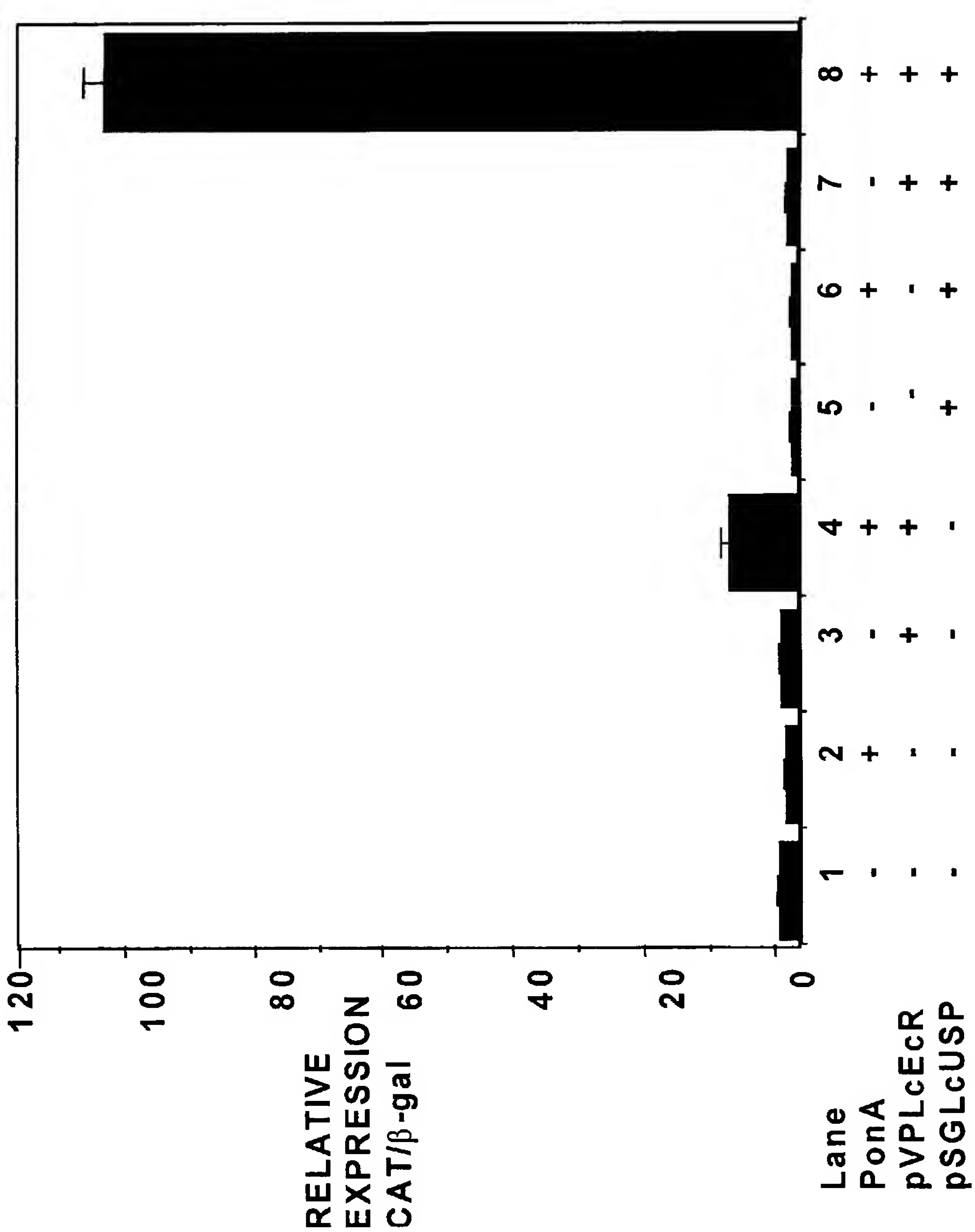

RELATIVE EXPRESSION CAT/β-gal
0
20
40
60
80
100
120
Lane 1 2 3 4 5 6 7 8
PonA - + - + - + - +
pVPLcEcR - - + + - - + +
pSGLcUSP - - - - + + + +

FIGURE 5

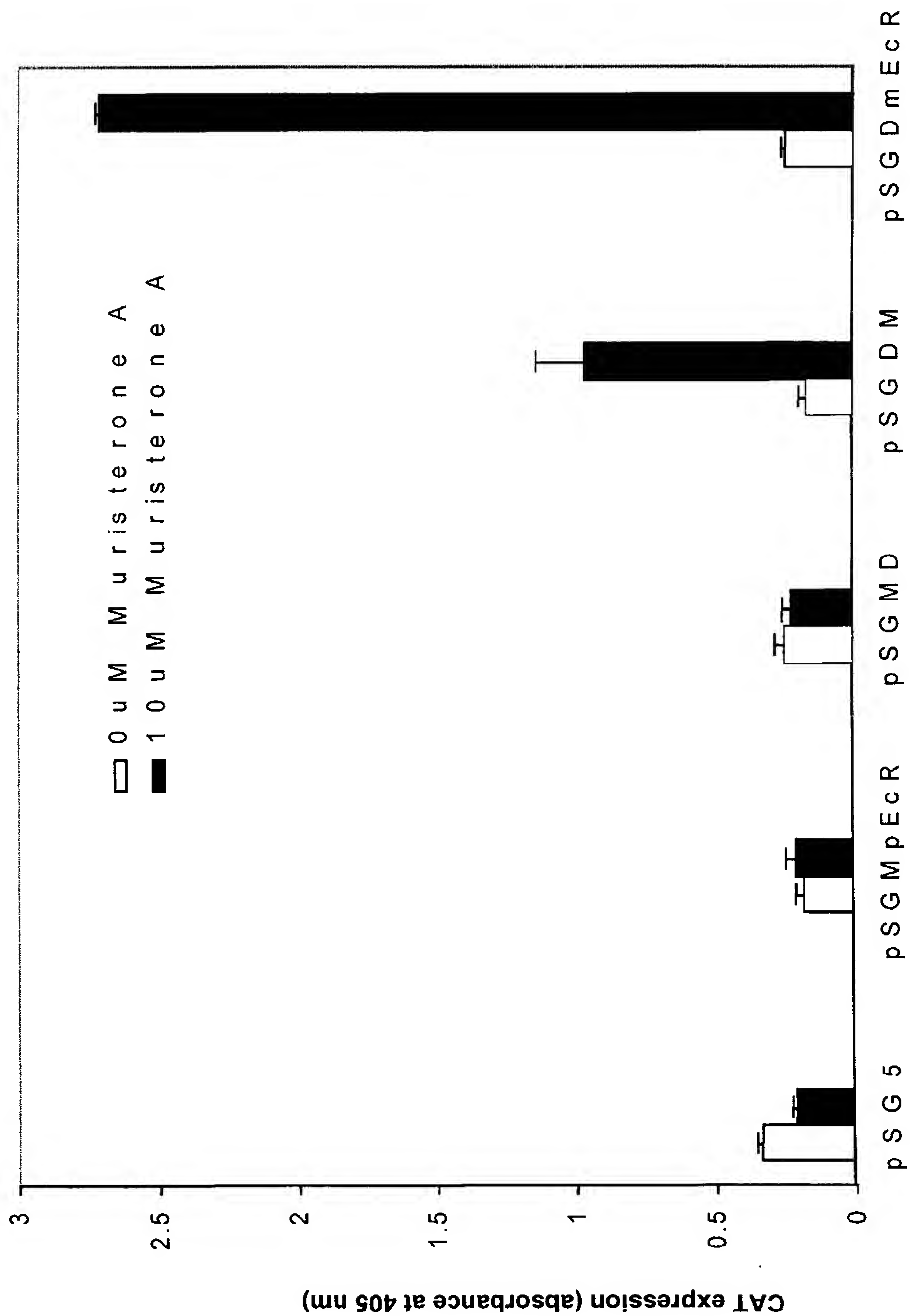

0 uM Muristerone A
10 uM Muristerone A
CAT expression (absorbance at 405 nm)
0
0.5
1
1.5
2
2.5
3
pSG5
pSGMpEcR
pSGMD
pSGDM
pSGDmEcR

FIGURE 6

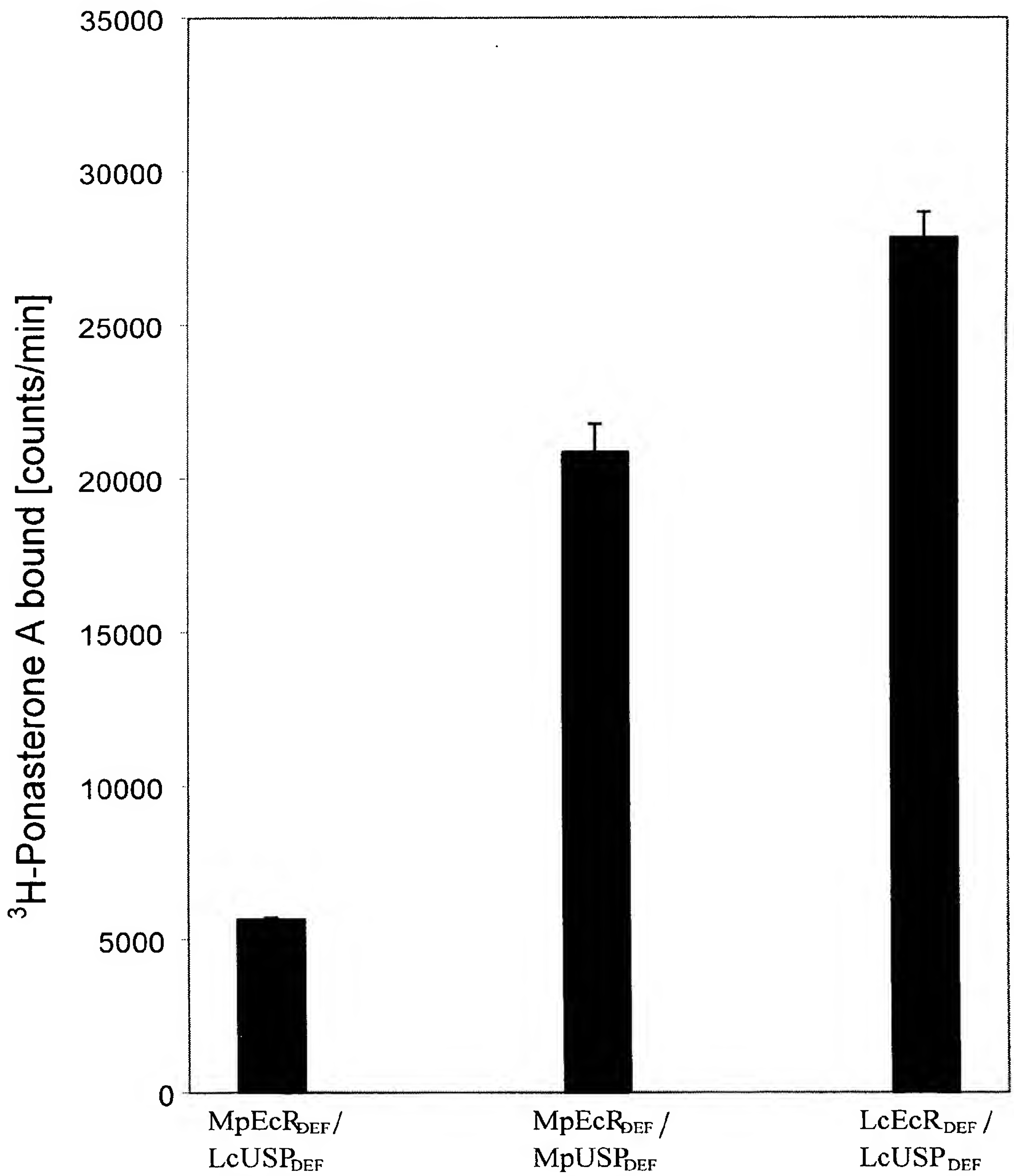
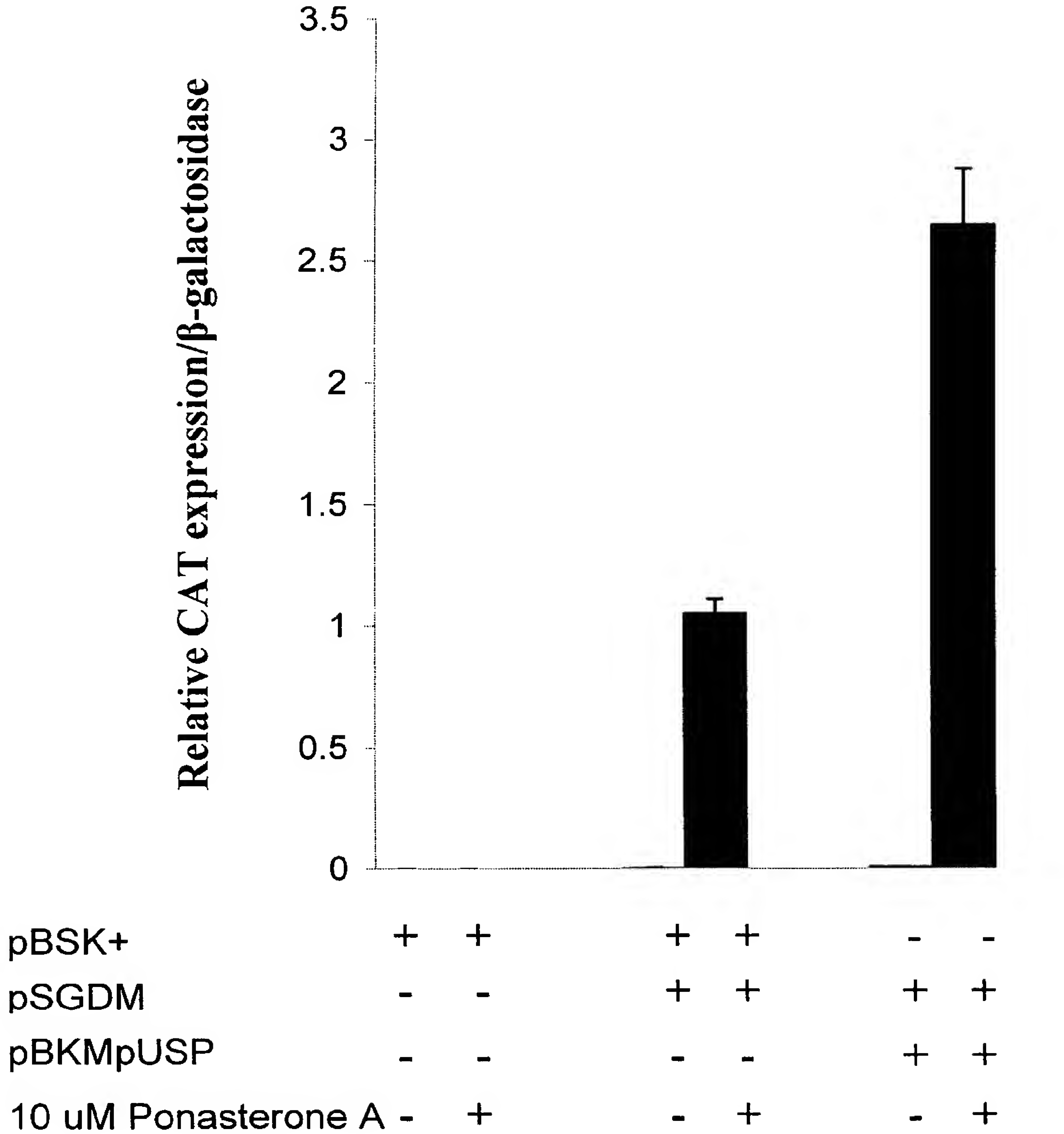

35000
30000
25000
20000
15000
10000
5000
0
3H-Ponasterone A bound [counts/min]
MpEcRDEF / LcUSPDEF
MpEcRDEF / MpUSPDEF
LcEcRDEF / LcUSPDEF

FIGURE 7



Relative CAT expression/β-galactosidase
3.5
3
2.5
2
1.5
1
0.5
0
pBSK+ + + + + - -
pSGDM - - + + + +
pBKMpUSP - - - - + +
10 uM Ponasterone A - + - + - +

FIGURE 8

SEQUENCE LISTING

<110> COMMONWEALTH SCIENTIFIC AND INDUSTRIAL RESEARCH ORGANISATION

<120> NOVEL GENETIC SEQUENCES ENCODING STEROID AND JUVENILE HORMONE RECEPTOR POLYPEPTIDES AND INSECTICIDAL MODALITIES THEREFOR II

<130> p:\oper\mro\ecdysone2.pct

<140> International application No. PCT/AU00/XXXXX
<141> 2000-06-30

<150> US 09/346470
<151> 1999-07-01

<160> 42
<170> PatentIn Ver. 2.0

<210> 1
<211> 2274
<212> DNA
<213> Lucilia cuprina

<220>
<221> CDS
<222> (1)..(2271)

<400> 1

```
atg atg aaa cga cgt tgg tct aat aat ggc ggt ttt gcc gct tta aaa    48
Met Met Lys Arg Arg Trp Ser Asn Asn Gly Gly Phe Ala Ala Leu Lys
  1               5                  10                  15

atg tta gaa gaa tcc tcc tca gaa gta acc tcc tcc tca aat ggt ctg    96
Met Leu Glu Glu Ser Ser Ser Glu Val Thr Ser Ser Ser Asn Gly Leu
             20                  25                  30

gtc ttg tca tcg gat ata aat atg tca cct tcc tcg ttg gat tca ccc   144
Val Leu Ser Ser Asp Ile Asn Met Ser Pro Ser Ser Leu Asp Ser Pro
         35                  40                  45

gtt tat ggc gat cag gaa atg tgg ctg tgt aac gat tca gct tca tat   192
Val Tyr Gly Asp Gln Glu Met Trp Leu Cys Asn Asp Ser Ala Ser Tyr
     50                  55                  60

aat aac agt cat cag cat agt gtt ata act tcg ctg cag ggc tgc acc   240
Asn Asn Ser His Gln His Ser Val Ile Thr Ser Leu Gln Gly Cys Thr
```

65 70 75 80

tca tca ttg ccg gcc caa aca acc att ata cct ctg tca gct tta ccc 288
Ser Ser Leu Pro Ala Gln Thr Thr Ile Ile Pro Leu Ser Ala Leu Pro
85 90 95

aat tcc aat aat gcc tcc ctg aat aat caa aat caa aat tat caa aat 336
Asn Ser Asn Asn Ala Ser Leu Asn Asn Gln Asn Gln Asn Tyr Gln Asn
100 105 110

ggt aat tcc atg aat aca aat tta tcg gtt aac aca aat aac agt gtt 384
Gly Asn Ser Met Asn Thr Asn Leu Ser Val Asn Thr Asn Asn Ser Val
115 120 125

gga gga ggt gga ggt ggt ggt ggt gta ccc ggt atg act tca ctc aat 432
Gly Gly Gly Gly Gly Gly Gly Gly Val Pro Gly Met Thr Ser Leu Asn
130 135 140

ggt ctg ggt ggt ggt ggt ggc agt caa gtg aat aat cac aat cac agc 480
Gly Leu Gly Gly Gly Gly Gly Ser Gln Val Asn Asn His Asn His Ser
145 150 155 160

cac aat cat tta cac cac aac agc aac agt aat cac agt aat agc agt 528
His Asn His Leu His His Asn Ser Asn Ser Asn His Ser Asn Ser Ser
165 170 175

tcc cac cac aca aat ggc cac atg ggt att ggc ggc ggt ggt ggt ggc 576
Ser His His Thr Asn Gly His Met Gly Ile Gly Gly Gly Gly Gly Gly
180 185 190

tta tcg gtc aat att aat ggt ccc aat atc gtt agc aat gcc caa cag 624
Leu Ser Val Asn Ile Asn Gly Pro Asn Ile Val Ser Asn Ala Gln Gln
195 200 205

tta aac tcg tta cag gcc tca caa aat ggc caa gtt att cat gcc aat 672
Leu Asn Ser Leu Gln Ala Ser Gln Asn Gly Gln Val Ile His Ala Asn
210 215 220

att ggc att cac agt atc atc agt aat gga tta aat cat cat cac cat 720
Ile Gly Ile His Ser Ile Ile Ser Asn Gly Leu Asn His His His His
225 230 235 240

cat cat atg aat aac agt agt atg atg cat cat aca ccc aga tct gaa 768
His His Met Asn Asn Ser Ser Met Met His His Thr Pro Arg Ser Glu
245 250 255

```
tca gct aat tcc ata tca tca ggt cgt gat gat ctt tca ccc tcg agc    816
Ser Ala Asn Ser Ile Ser Ser Gly Arg Asp Asp Leu Ser Pro Ser Ser
            260                 265                 270

agt ctt aat ggc ttc tca aca agc gat gct agt gat gtt aag aaa atc    864
Ser Leu Asn Gly Phe Ser Thr Ser Asp Ala Ser Asp Val Lys Lys Ile
        275                 280                 285

aaa aaa ggt cct gcg ccc cgt tta caa gag gaa ctg tgt ctg gtg tgt    912
Lys Lys Gly Pro Ala Pro Arg Leu Gln Glu Glu Leu Cys Leu Val Cys
    290                 295                 300

ggt gat cgg gcg tcc ggt tat cat tat aac gca ctc acc tgt gaa ggc    960
Gly Asp Arg Ala Ser Gly Tyr His Tyr Asn Ala Leu Thr Cys Glu Gly
305                 310                 315                 320

tgt aag ggg ttc ttt cga cgg agt gtt acc aaa aat gcg gtg tat tgt    1008
Cys Lys Gly Phe Phe Arg Arg Ser Val Thr Lys Asn Ala Val Tyr Cys
                325                 330                 335

tgt aaa ttt ggt cat gcc tgc gaa atg gac atg tat atg cga cgt aaa    1056
Cys Lys Phe Gly His Ala Cys Glu Met Asp Met Tyr Met Arg Arg Lys
            340                 345                 350

tgt cag gaa tgt agg ctg aaa aaa tgt ttg gct gtg ggc atg cgg ccg    1104
Cys Gln Glu Cys Arg Leu Lys Lys Cys Leu Ala Val Gly Met Arg Pro
        355                 360                 365

gaa tgt gtg gtg ccc gaa aac cag tgt gca atg aaa cga cgc gaa aag    1152
Glu Cys Val Val Pro Glu Asn Gln Cys Ala Met Lys Arg Arg Glu Lys
    370                 375                 380

aaa gca caa aaa gag aag gat aaa ata cag acc agt gtg tgt gca acg    1200
Lys Ala Gln Lys Glu Lys Asp Lys Ile Gln Thr Ser Val Cys Ala Thr
385                 390                 395                 400

gaa att aaa aag gaa ata ctc gat tta atg aca tgt gaa ccg cca tca    1248
Glu Ile Lys Lys Glu Ile Leu Asp Leu Met Thr Cys Glu Pro Pro Ser
                405                 410                 415

cat cca acg tgt ccg ctg tta cct gaa gac att ttg gct aaa tgt caa    1296
His Pro Thr Cys Pro Leu Leu Pro Glu Asp Ile Leu Ala Lys Cys Gln
            420                 425                 430

gct cgt aat ata cct cct tta tcg tac aat caa ttg gca gtt ata tat    1344
Ala Arg Asn Ile Pro Pro Leu Ser Tyr Asn Gln Leu Ala Val Ile Tyr
```

435 440 445

aaa tta ata tgg tat caa gat ggc tac gaa cag cca tcc gag gaa gat 1392
Lys Leu Ile Trp Tyr Gln Asp Gly Tyr Glu Gln Pro Ser Glu Glu Asp
450 455 460

ctc aaa cgt ata atg agt tca ccc gat gaa aat gaa agt caa cac gat 1440
Leu Lys Arg Ile Met Ser Ser Pro Asp Glu Asn Glu Ser Gln His Asp
465 470 475 480

gca tca ttt cgt cat ata aca gaa atc act ata cta aca gta caa tta 1488
Ala Ser Phe Arg His Ile Thr Glu Ile Thr Ile Leu Thr Val Gln Leu
485 490 495

att gtg gaa ttt gcc aag ggt ttg cca gcg ttt acc aaa ata cca caa 1536
Ile Val Glu Phe Ala Lys Gly Leu Pro Ala Phe Thr Lys Ile Pro Gln
500 505 510

gag gat caa ata aca cta tta aag gcc tgc tca tca gaa gtt atg atg 1584
Glu Asp Gln Ile Thr Leu Leu Lys Ala Cys Ser Ser Glu Val Met Met
515 520 525

ttg cga atg gca cga cgt tac gat cac aat tca gat tcg ata ttc ttt 1632
Leu Arg Met Ala Arg Arg Tyr Asp His Asn Ser Asp Ser Ile Phe Phe
530 535 540

cc aat aat cga tcg tat acg cgt gac tct tat aaa atg gct ggc atg 1680
Ala Asn Asn Arg Ser Tyr Thr Arg Asp Ser Tyr Lys Met Ala Gly Met
545 550 555 560

gct gat aat att gag gat ctg ctg cat ttc tgt cga caa atg tac tcg 1728
Ala Asp Asn Ile Glu Asp Leu Leu His Phe Cys Arg Gln Met Tyr Ser
565 570 575

atg aaa gtg gac aat gtc gaa tat gct cta ctc act gcc att gtg atc 1776
Met Lys Val Asp Asn Val Glu Tyr Ala Leu Leu Thr Ala Ile Val Ile
580 585 590

ttt tcc gat cgg ccg ggt ctc gaa gaa gcc gaa cta gtc gaa gcg ata 1824
Phe Ser Asp Arg Pro Gly Leu Glu Glu Ala Glu Leu Val Glu Ala Ile
595 600 605

caa agt tac tac atc gat aca ctc cgc att tac ata ctt aat cgc cat 1872
Gln Ser Tyr Tyr Ile Asp Thr Leu Arg Ile Tyr Ile Leu Asn Arg His
610 615 620

```
tgc ggc gat ccc atg agt ctc gta ttc ttt gcc aag ctt ctg tca att    1920
Cys Gly Asp Pro Met Ser Leu Val Phe Phe Ala Lys Leu Leu Ser Ile
625                 630                 635                 640

cta acc gaa ctg cgt acg ttg ggc aat caa aat gcc gaa atg tgt ttc    1968
Leu Thr Glu Leu Arg Thr Leu Gly Asn Gln Asn Ala Glu Met Cys Phe
                645                 650                 655

tcg ttg aaa ttg aaa aat cgc aaa ctg cca aaa ttc ctc gaa gag atc    2016
Ser Leu Lys Leu Lys Asn Arg Lys Leu Pro Lys Phe Leu Glu Glu Ile
            660                 665                 670

tgg gat gta cat gcc att cca ccc tca gtg cag tca cac ata cag gct    2064
Trp Asp Val His Ala Ile Pro Pro Ser Val Gln Ser His Ile Gln Ala
        675                 680                 685

acc cag gcg gaa aag gcc gcc cag gaa gct cag gca aca aca tcg gcc    2112
Thr Gln Ala Glu Lys Ala Ala Gln Glu Ala Gln Ala Thr Thr Ser Ala
    690                 695                 700

att tca gca gcc gcc acc tca tct tcc tcc ata aat acc tcg atg gca    2160
Ile Ser Ala Ala Ala Thr Ser Ser Ser Ser Ile Asn Thr Ser Met Ala
705                 710                 715                 720

aca tca tcc tca tca tcg tta tcg cca tcg gcg gcc tca aca ccc aat    2208
Thr Ser Ser Ser Ser Ser Leu Ser Pro Ser Ala Ala Ser Thr Pro Asn
                725                 730                 735

ggt ggt gcc gtc gat tat gtt ggc acc gat atg agt atg agt tta gta    2256
Gly Gly Ala Val Asp Tyr Val Gly Thr Asp Met Ser Met Ser Leu Val
            740                 745                 750

caa tcg gat aat gca tag                                            2274
Gln Ser Asp Asn Ala
        755


<210> 2
<211> 757
<212> PRT
<213> Lucilia cuprina

<400> 2
Met Met Lys Arg Arg Trp Ser Asn Asn Gly Gly Phe Ala Ala Leu Lys
  1               5                  10                  15
```

```
Met Leu Glu Glu Ser Ser Ser Glu Val Thr Ser Ser Ser Asn Gly Leu
            20                  25                  30

Val Leu Ser Ser Asp Ile Asn Met Ser Pro Ser Ser Leu Asp Ser Pro
        35                  40                  45

Val Tyr Gly Asp Gln Glu Met Trp Leu Cys Asn Asp Ser Ala Ser Tyr
    50                  55                  60

Asn Asn Ser His Gln His Ser Val Ile Thr Ser Leu Gln Gly Cys Thr
65                  70                  75                  80

Ser Ser Leu Pro Ala Gln Thr Thr Ile Ile Pro Leu Ser Ala Leu Pro
                85                  90                  95

Asn Ser Asn Asn Ala Ser Leu Asn Asn Gln Asn Gln Asn Tyr Gln Asn
            100                 105                 110

Gly Asn Ser Met Asn Thr Asn Leu Ser Val Asn Thr Asn Asn Ser Val
        115                 120                 125

Gly Gly Gly Gly Gly Gly Gly Gly Val Pro Gly Met Thr Ser Leu Asn
    130                 135                 140

Gly Leu Gly Gly Gly Gly Gly Ser Gln Val Asn Asn His Asn His Ser
145                 150                 155                 160

His Asn His Leu His His Asn Ser Asn Ser Asn His Ser Asn Ser Ser
                165                 170                 175

Ser His His Thr Asn Gly His Met Gly Ile Gly Gly Gly Gly Gly Gly
            180                 185                 190

Leu Ser Val Asn Ile Asn Gly Pro Asn Ile Val Ser Asn Ala Gln Gln
        195                 200                 205

Leu Asn Ser Leu Gln Ala Ser Gln Asn Gly Gln Val Ile His Ala Asn
    210                 215                 220

Ile Gly Ile His Ser Ile Ile Ser Asn Gly Leu Asn His His His His
225                 230                 235                 240

His His Met Asn Asn Ser Ser Met Met His His Thr Pro Arg Ser Glu
                245                 250                 255

Ser Ala Asn Ser Ile Ser Ser Gly Arg Asp Asp Leu Ser Pro Ser Ser
```

260 265 270

Ser Leu Asn Gly Phe Ser Thr Ser Asp Ala Ser Asp Val Lys Lys Ile
275 280 285

Lys Lys Gly Pro Ala Pro Arg Leu Gln Glu Glu Leu Cys Leu Val Cys
290 295 300

Gly Asp Arg Ala Ser Gly Tyr His Tyr Asn Ala Leu Thr Cys Glu Gly
305 310 315 320

Cys Lys Gly Phe Phe Arg Arg Ser Val Thr Lys Asn Ala Val Tyr Cys
325 330 335

Cys Lys Phe Gly His Ala Cys Glu Met Asp Met Tyr Met Arg Arg Lys
340 345 350

Cys Gln Glu Cys Arg Leu Lys Lys Cys Leu Ala Val Gly Met Arg Pro
355 360 365

Glu Cys Val Val Pro Glu Asn Gln Cys Ala Met Lys Arg Arg Glu Lys
370 375 380

Lys Ala Gln Lys Glu Lys Asp Lys Ile Gln Thr Ser Val Cys Ala Thr
385 390 395 400

Glu Ile Lys Lys Glu Ile Leu Asp Leu Met Thr Cys Glu Pro Pro Ser
405 410 415

His Pro Thr Cys Pro Leu Leu Pro Glu Asp Ile Leu Ala Lys Cys Gln
420 425 430

Ala Arg Asn Ile Pro Pro Leu Ser Tyr Asn Gln Leu Ala Val Ile Tyr
435 440 445

Lys Leu Ile Trp Tyr Gln Asp Gly Tyr Glu Gln Pro Ser Glu Glu Asp
450 455 460

Leu Lys Arg Ile Met Ser Ser Pro Asp Glu Asn Glu Ser Gln His Asp
465 470 475 480

Ala Ser Phe Arg His Ile Thr Glu Ile Thr Ile Leu Thr Val Gln Leu
485 490 495

Ile Val Glu Phe Ala Lys Gly Leu Pro Ala Phe Thr Lys Ile Pro Gln

500 505 510

Glu Asp Gln Ile Thr Leu Leu Lys Ala Cys Ser Ser Glu Val Met Met
515 520 525

Leu Arg Met Ala Arg Arg Tyr Asp His Asn Ser Asp Ser Ile Phe Phe
530 535 540

Ala Asn Asn Arg Ser Tyr Thr Arg Asp Ser Tyr Lys Met Ala Gly Met
545 550 555 560

Ala Asp Asn Ile Glu Asp Leu Leu His Phe Cys Arg Gln Met Tyr Ser
565 570 575

Met Lys Val Asp Asn Val Glu Tyr Ala Leu Leu Thr Ala Ile Val Ile
580 585 590

Phe Ser Asp Arg Pro Gly Leu Glu Glu Ala Glu Leu Val Glu Ala Ile
595 600 605

Gln Ser Tyr Tyr Ile Asp Thr Leu Arg Ile Tyr Ile Leu Asn Arg His
610 615 620

Cys Gly Asp Pro Met Ser Leu Val Phe Phe Ala Lys Leu Leu Ser Ile
625 630 635 640

Leu Thr Glu Leu Arg Thr Leu Gly Asn Gln Asn Ala Glu Met Cys Phe
645 650 655

Ser Leu Lys Leu Lys Asn Arg Lys Leu Pro Lys Phe Leu Glu Glu Ile
660 665 670

Trp Asp Val His Ala Ile Pro Pro Ser Val Gln Ser His Ile Gln Ala
675 680 685

Thr Gln Ala Glu Lys Ala Ala Gln Glu Ala Gln Ala Thr Thr Ser Ala
690 695 700

Ile Ser Ala Ala Ala Thr Ser Ser Ser Ser Ile Asn Thr Ser Met Ala
705 710 715 720

Thr Ser Ser Ser Ser Ser Leu Ser Pro Ser Ala Ala Ser Thr Pro Asn
725 730 735

Gly Gly Ala Val Asp Tyr Val Gly Thr Asp Met Ser Met Ser Leu Val
740 745 750

```
Gln Ser Asp Asn Ala
        755


<210> 3
<211> 2453
<212> DNA
<213> Lucilia cuprina

<220>
<221> CDS
<222> (17)..(1417)

<400> 3
aattcggcac gagaaa atg gat aac ggc gag caa gat gct ggg ttc cga ttg 52
                  Met Asp Asn Gly Glu Gln Asp Ala Gly Phe Arg Leu
                  1               5                   10

gca ccg atg tct ccg cag gag ata aag cca gac att tca cta ctc aat    100
Ala Pro Met Ser Pro Gln Glu Ile Lys Pro Asp Ile Ser Leu Leu Asn
        15                  20                  25

gaa aat aat acg agt agt tat tcg ccc aaa cct gga agt cct aat cca    148
Glu Asn Asn Thr Ser Ser Tyr Ser Pro Lys Pro Gly Ser Pro Asn Pro
    30                  35                  40

ttt gcc atc gga ttg cag gca ata aat gca gtc gct gcc gcg aat gcc    196
Phe Ala Ile Gly Leu Gln Ala Ile Asn Ala Val Ala Ala Ala Asn Ala
45                  50                  55                  60

aat aac caa aat caa atg ttg caa act acg cca cca caa cag cag cag    244
Asn Asn Gln Asn Gln Met Leu Gln Thr Thr Pro Pro Gln Gln Gln Gln
                65                  70                  75

tat cca cca aat cac ccc ctt agt ggt tcg aaa cac ttg tgt tcc att    292
Tyr Pro Pro Asn His Pro Leu Ser Gly Ser Lys His Leu Cys Ser Ile
            80                  85                  90

tgt gga gac cgc gcc agt gga aaa cat tat ggg gtc tac agt tgt gag    340
Cys Gly Asp Arg Ala Ser Gly Lys His Tyr Gly Val Tyr Ser Cys Glu
        95                  100                 105

ggt tgt aaa ggg ttc ttc aaa cgt acc gta cgc aag gac ttg aca tat    388
Gly Cys Lys Gly Phe Phe Lys Arg Thr Val Arg Lys Asp Leu Thr Tyr
    110                 115                 120
```

```
gct tgt cgt gag gac aga aat tgc att ata gat aaa cga caa aga aat    436
Ala Cys Arg Glu Asp Arg Asn Cys Ile Ile Asp Lys Arg Gln Arg Asn
125                 130                 135                 140

cgt tgc cag tat tgt cgt tat caa aag tgt tta gct tgt ggc atg aaa    484
Arg Cys Gln Tyr Cys Arg Tyr Gln Lys Cys Leu Ala Cys Gly Met Lys
                145                 150                 155

cgc gaa gcg gtc caa gag gaa cga caa cgt ggt act cgt gct gct aac    532
Arg Glu Ala Val Gln Glu Glu Arg Gln Arg Gly Thr Arg Ala Ala Asn
            160                 165                 170

gct aga gct gct ggt gct ggc ggt ggt gga gga ggt ggt ggt ggg gta    580
Ala Arg Ala Ala Gly Ala Gly Gly Gly Gly Gly Gly Gly Gly Gly Val
        175                 180                 185

agc aat gtg gtt ggt gct ggc gga gaa gac ttt aaa ccc agc agt tca    628
Ser Asn Val Val Gly Ala Gly Gly Glu Asp Phe Lys Pro Ser Ser Ser
    190                 195                 200

tta cgt gat ctc act ata gaa cgc atc att gaa gcc gag caa aag gct    676
Leu Arg Asp Leu Thr Ile Glu Arg Ile Ile Glu Ala Glu Gln Lys Ala
205                 210                 215                 220

gaa tct ttg agc ggt gat aac gtg ttg ccc ttt ttg cgc gtt ggc aac    724
Glu Ser Leu Ser Gly Asp Asn Val Leu Pro Phe Leu Arg Val Gly Asn
                225                 230                 235

aat tcc atg gta caa cac gac tac aaa ggc gcg gta tct cat ctc tgc    772
Asn Ser Met Val Gln His Asp Tyr Lys Gly Ala Val Ser His Leu Cys
            240                 245                 250

cag atg gtt aac aaa caa ctc tac caa atg gtt gaa tat gca cgt cga    820
Gln Met Val Asn Lys Gln Leu Tyr Gln Met Val Glu Tyr Ala Arg Arg
        255                 260                 265

aca cca cat ttt aca cat ttg cag cgt gag gat cag ata cta ttg tta    868
Thr Pro His Phe Thr His Leu Gln Arg Glu Asp Gln Ile Leu Leu Leu
    270                 275                 280

aag gct ggc tgg aat gaa ctg cta att gca aat gtt gcc tgg tgc agt    916
Lys Ala Gly Trp Asn Glu Leu Leu Ile Ala Asn Val Ala Trp Cys Ser
285                 290                 295                 300

att gag tct ctg gat gcc gaa tat gcc tct cct ggt acg gta cat gac    964
Ile Glu Ser Leu Asp Ala Glu Tyr Ala Ser Pro Gly Thr Val His Asp
```

```
                305                 310                 315

ggt tct ttt ggt cgg cgt tca cca gtg cgt cag ccc caa caa ctc ttc    1012
Gly Ser Phe Gly Arg Arg Ser Pro Val Arg Gln Pro Gln Gln Leu Phe
            320                 325                 330

ctt aat cag aat ttc tcg tat cat cgc aat agt gct att aag gcc aat    1060
Leu Asn Gln Asn Phe Ser Tyr His Arg Asn Ser Ala Ile Lys Ala Asn
        335                 340                 345

gtt gtt tca att ttc gat cgt atc ctc tcg gag ttg agc atc aaa atg    1108
Val Val Ser Ile Phe Asp Arg Ile Leu Ser Glu Leu Ser Ile Lys Met
    350                 355                 360

aaa cgt ctt aac atc gat cgc tcg gag ttg tcg tgt ctg aag gca atc    1156
Lys Arg Leu Asn Ile Asp Arg Ser Glu Leu Ser Cys Leu Lys Ala Ile
365                 370                 375                 380

ata ctc ttc aat cca gac ata cgc ggt ctg aaa tgt cga gcc gac gtc    1204
Ile Leu Phe Asn Pro Asp Ile Arg Gly Leu Lys Cys Arg Ala Asp Val
                385                 390                 395

gag gta tgt cgt gaa aaa atc tat gcc tgt ctg gac gaa cac tgc cgc    1252
Glu Val Cys Arg Glu Lys Ile Tyr Ala Cys Leu Asp Glu His Cys Arg
            400                 405                 410

aca gaa cat cca ggt gat gat ggc cgc ttt gct cag cta cta cta agg    1300
Thr Glu His Pro Gly Asp Asp Gly Arg Phe Ala Gln Leu Leu Leu Arg
        415                 420                 425

ttg ccc gca ttg cgt tcc atc agt ctc aaa tgt ctc gat cat ttg ttt    1348
Leu Pro Ala Leu Arg Ser Ile Ser Leu Lys Cys Leu Asp His Leu Phe
    430                 435                 440

ttc ttc cgt tta ata ggc gaa aga gca ttg gag gaa tta att gct gag    1396
Phe Phe Arg Leu Ile Gly Glu Arg Ala Leu Glu Glu Leu Ile Ala Glu
445                 450                 455                 460

caa ttg gaa gct cct atc tgc taagaaattt gaaagttgta ctaaaataaa       1447
Gln Leu Glu Ala Pro Ile Cys
                465

acacaacatc caaaggactg tgttgtgaaa tgaatgatga tagagaaatt atttgttggt 1507

gcttcaaaga atcaatcgtt aaattaaaag gtgatcataa aggccaagcc tgggaagcat 1567
```

catataacac aaatggatac gatatttgaa tagtagcaat agttatattc agtctttttt 1627

atttttttctt tttttttttt tttgtcaaca ttaatattct tgctctttat gtagattaat 1687

gaaaaacaaa acagaaaaca caacatacac aaagacacat ccatatacat ccaatactga 1747

tagcaaaaat aattgttgaa ttgatggcct tgttggtggt ggctgtgtat ctaagaattt 1807

attccctaaa atttggactt acatttaatt attaagctca taatatgtga ttgagctctt 1867

gctgcatttc cttaaccttt ccacacagtc tccaaaacag tctgttaatg gcgtgaagtg 1927

ataatgatct taagttattt tgcaaaccaa acaaaacgga tgttgaaaaa caaatatata 1987

atttataata atctctacaa atgactttat ttcaataatg acagctgata tttttatatg 2047

tacatgtatt ttagttaaaa tatgccttta aagtgtaaat catatagatt ttacttatat 2107

attatttaaa catttttat gtttatattg ttttagatct gcctttttc aaaattcaat 2167

acgaatgaat ttaaaagttc aatactttag ttaaataatt tgacaaaaat acttattaca 2227

atatcacgtt taggtttttt ttactccttt actttagtta taaacatgtt tttattgtgt 2287

attatataaa tgaaaactga catacatact aacaaacact tatacataga tatattatac 2347

atacaaacat attaatatat tattaaatat agatatattt tattctgttg aaacaacaaa 2407

aaaaaaaaaa aaaactcgag actagttctc tccctcgtgc cgaatt 2453

<210> 4
<211> 467
<212> PRT
<213> Lucilia cuprina

<400> 4
Met Asp Asn Gly Glu Gln Asp Ala Gly Phe Arg Leu Ala Pro Met Ser
1 5 10 15

Pro Gln Glu Ile Lys Pro Asp Ile Ser Leu Leu Asn Glu Asn Asn Thr
20 25 30

Ser Ser Tyr Ser Pro Lys Pro Gly Ser Pro Asn Pro Phe Ala Ile Gly
35 40 45

```
Leu Gln Ala Ile Asn Ala Val Ala Ala Ala Asn Ala Asn Asn Gln Asn
     50                  55                  60

Gln Met Leu Gln Thr Thr Pro Pro Gln Gln Gln Gln Tyr Pro Pro Asn
 65                  70                  75                  80

His Pro Leu Ser Gly Ser Lys His Leu Cys Ser Ile Cys Gly Asp Arg
                 85                  90                  95

Ala Ser Gly Lys His Tyr Gly Val Tyr Ser Cys Glu Gly Cys Lys Gly
            100                 105                 110

Phe Phe Lys Arg Thr Val Arg Lys Asp Leu Thr Tyr Ala Cys Arg Glu
        115                 120                 125

Asp Arg Asn Cys Ile Ile Asp Lys Arg Gln Arg Asn Arg Cys Gln Tyr
    130                 135                 140

Cys Arg Tyr Gln Lys Cys Leu Ala Cys Gly Met Lys Arg Glu Ala Val
145                 150                 155                 160

Gln Glu Glu Arg Gln Arg Gly Thr Arg Ala Ala Asn Ala Arg Ala Ala
                165                 170                 175

Gly Ala Gly Gly Gly Gly Gly Gly Gly Gly Gly Val Ser Asn Val Val
            180                 185                 190

Gly Ala Gly Gly Glu Asp Phe Lys Pro Ser Ser Ser Leu Arg Asp Leu
        195                 200                 205

Thr Ile Glu Arg Ile Ile Glu Ala Glu Gln Lys Ala Glu Ser Leu Ser
    210                 215                 220

Gly Asp Asn Val Leu Pro Phe Leu Arg Val Gly Asn Asn Ser Met Val
225                 230                 235                 240

Gln His Asp Tyr Lys Gly Ala Val Ser His Leu Cys Gln Met Val Asn
                245                 250                 255

Lys Gln Leu Tyr Gln Met Val Glu Tyr Ala Arg Arg Thr Pro His Phe
            260                 265                 270

Thr His Leu Gln Arg Glu Asp Gln Ile Leu Leu Leu Lys Ala Gly Trp
        275                 280                 285

Asn Glu Leu Leu Ile Ala Asn Val Ala Trp Cys Ser Ile Glu Ser Leu
```

290 295 300

Asp Ala Glu Tyr Ala Ser Pro Gly Thr Val His Asp Gly Ser Phe Gly
305 310 315 320

Arg Arg Ser Pro Val Arg Gln Pro Gln Gln Leu Phe Leu Asn Gln Asn
325 330 335

Phe Ser Tyr His Arg Asn Ser Ala Ile Lys Ala Asn Val Val Ser Ile
340 345 350

Phe Asp Arg Ile Leu Ser Glu Leu Ser Ile Lys Met Lys Arg Leu Asn
355 360 365

Ile Asp Arg Ser Glu Leu Ser Cys Leu Lys Ala Ile Ile Leu Phe Asn
370 375 380

Pro Asp Ile Arg Gly Leu Lys Cys Arg Ala Asp Val Glu Val Cys Arg
385 390 395 400

Glu Lys Ile Tyr Ala Cys Leu Asp Glu His Cys Arg Thr Glu His Pro
405 410 415

Gly Asp Asp Gly Arg Phe Ala Gln Leu Leu Leu Arg Leu Pro Ala Leu
420 425 430

Arg Ser Ile Ser Leu Lys Cys Leu Asp His Leu Phe Phe Phe Arg Leu
435 440 445

Ile Gly Glu Arg Ala Leu Glu Glu Leu Ile Ala Glu Gln Leu Glu Ala
450 455 460

Pro Ile Cys
465

<210> 5
<211> 1596
<212> DNA
<213> Lucilia cuprina

<220>
<221> CDS
<222> (193)..(1593)

<400> 5
aattcggcac gagtcgacgc gaaaactttt cattcattag tcaaccagaa tatagacatt 60

```
ctttgtttgt aaaaaaatct tgtaataatt aaaaatcaat tatcaaaact tatagttaaa 120

tgtattaaat aaagattgtg tgtgacagaa acaaattagt gagatctctt gatacgggaa 180

aatataatca aa atg gat aac ggc gag caa gat gct ggg ttc cga ttg gca 231
              Met Asp Asn Gly Glu Gln Asp Ala Gly Phe Arg Leu Ala
                1               5                   10

ccg atg tct ccg cag gag ata aag cca gac att tca cta ctc aat gaa   279
Pro Met Ser Pro Gln Glu Ile Lys Pro Asp Ile Ser Leu Leu Asn Glu
     15                  20                  25

aat aat acg agt agt tat tcg ccc aaa cct gga agt cct aat cca ttt   327
Asn Asn Thr Ser Ser Tyr Ser Pro Lys Pro Gly Ser Pro Asn Pro Phe
 30                  35                  40                  45

gcc atc gga ttg cag gca ata aat gca gtc gct gcc gcg aat gcc aat   375
Ala Ile Gly Leu Gln Ala Ile Asn Ala Val Ala Ala Ala Asn Ala Asn
                 50                  55                  60

aac caa aat caa atg ttg caa act acg cca cca caa cag cag cag tat   423
Asn Gln Asn Gln Met Leu Gln Thr Thr Pro Pro Gln Gln Gln Gln Tyr
             65                  70                  75

cca cca aat cac ccc ctt agt ggt tcg aaa cac ttg tgt tcc att tgt   471
Pro Pro Asn His Pro Leu Ser Gly Ser Lys His Leu Cys Ser Ile Cys
         80                  85                  90

gga gac cgc gcc agt gga aaa cat tat ggg gtc tac agt tgt gag ggt   519
Gly Asp Arg Ala Ser Gly Lys His Tyr Gly Val Tyr Ser Cys Glu Gly
     95                 100                 105

tgt aaa ggg ttc ttc aaa cgt acc gta cgc aag gac ttg aca tat gct   567
Cys Lys Gly Phe Phe Lys Arg Thr Val Arg Lys Asp Leu Thr Tyr Ala
110                 115                 120                 125

tgt cgt gag gac aga aat tgc att ata gat aaa cga caa aga aat cgt   615
Cys Arg Glu Asp Arg Asn Cys Ile Ile Asp Lys Arg Gln Arg Asn Arg
                130                 135                 140

tgc cag tat tgt cgt tat caa aag tgt tta gct tgt ggc atg aaa cgc   663
Cys Gln Tyr Cys Arg Tyr Gln Lys Cys Leu Ala Cys Gly Met Lys Arg
            145                 150                 155

gaa gcg gtc caa gag gaa cga caa cgt ggt act cgt gct gct aac gct   711
```

```
Glu Ala Val Gln Glu Glu Arg Gln Arg Gly Thr Arg Ala Ala Asn Ala
        160                 165                 170

aga gct gct ggt gct ggc ggt ggt gga gga ggt ggt ggt ggg gta agc     759
Arg Ala Ala Gly Ala Gly Gly Gly Gly Gly Gly Gly Gly Gly Val Ser
    175                 180                 185

aat gtg gtt ggt gct ggc gga gaa gac ttt aaa ccc agc agt tca tta     807
Asn Val Val Gly Ala Gly Gly Glu Asp Phe Lys Pro Ser Ser Ser Leu
190                 195                 200                 205

cgt gat ctc act ata gaa cgc atc att gaa gcc gag caa aag gct gaa     855
Arg Asp Leu Thr Ile Glu Arg Ile Ile Glu Ala Glu Gln Lys Ala Glu
                210                 215                 220

tct ttg agc ggt gat aac gtg ttg ccc ttt ttg cgc gtt ggc aac aat     903
Ser Leu Ser Gly Asp Asn Val Leu Pro Phe Leu Arg Val Gly Asn Asn
            225                 230                 235

tcc atg gta caa cac gac tac aaa ggc gcg gta tct cat ctc tgc cag     951
Ser Met Val Gln His Asp Tyr Lys Gly Ala Val Ser His Leu Cys Gln
        240                 245                 250

atg gtt aac aaa caa ctc tac caa atg gtt gaa tat gca cgt cga aca     999
Met Val Asn Lys Gln Leu Tyr Gln Met Val Glu Tyr Ala Arg Arg Thr
    255                 260                 265

cca cat ttt aca cat ttg cag cgt gag gat cag ata cta ttg tta aag     1047
Pro His Phe Thr His Leu Gln Arg Glu Asp Gln Ile Leu Leu Leu Lys
270                 275                 280                 285

gct ggc tgg aat gaa ctg cta att gca aat gtt gcc tgg tgc agt att     1095
Ala Gly Trp Asn Glu Leu Leu Ile Ala Asn Val Ala Trp Cys Ser Ile
                290                 295                 300

gag tct ctg gat gcc gaa tat gcc tct cct ggt acg gta cat gac ggt     1143
Glu Ser Leu Asp Ala Glu Tyr Ala Ser Pro Gly Thr Val His Asp Gly
            305                 310                 315

tct ttt ggt cgg cgt tca cca gtg cgt cag ccc caa caa ctc ttc ctt     1191
Ser Phe Gly Arg Arg Ser Pro Val Arg Gln Pro Gln Gln Leu Phe Leu
        320                 325                 330

aat cag aat ttc tcg tat cat cgc aat agt gct att aag gcc aat gtt     1239
Asn Gln Asn Phe Ser Tyr His Arg Asn Ser Ala Ile Lys Ala Asn Val
    335                 340                 345
```

```
gtt tca att ttc gat cgt atc ctc tcg gag ttg agc atc aaa atg aaa    1287
Val Ser Ile Phe Asp Arg Ile Leu Ser Glu Leu Ser Ile Lys Met Lys
350                 355                 360                 365

cgt ctt aac atc gat cgc tcg gag ttg tcg tgt ctg aag gca atc ata    1335
Arg Leu Asn Ile Asp Arg Ser Glu Leu Ser Cys Leu Lys Ala Ile Ile
                370                 375                 380

ctc ttc aat cca gac ata cgc ggt ctg aaa tgt cga gcc gac gtc gag    1383
Leu Phe Asn Pro Asp Ile Arg Gly Leu Lys Cys Arg Ala Asp Val Glu
            385                 390                 395

gta tgt cgt gaa aaa atc tat gcc tgt ctg gac gaa cac tgc cgc aca    1431
Val Cys Arg Glu Lys Ile Tyr Ala Cys Leu Asp Glu His Cys Arg Thr
        400                 405                 410

gaa cat cca ggt gat gat ggc cgc ttt gct cag cta cta cta agg ttg    1479
Glu His Pro Gly Asp Asp Gly Arg Phe Ala Gln Leu Leu Leu Arg Leu
    415                 420                 425

ccc gca ttg cgt tcc atc agt ctc aaa tgt ctc gat cat ttg ttt ttc    1527
Pro Ala Leu Arg Ser Ile Ser Leu Lys Cys Leu Asp His Leu Phe Phe
430                 435                 440                 445

ttc cgt tta ata ggc gaa aga gca ttg gag gaa tta att gct gag caa    1575
Phe Arg Leu Ile Gly Glu Arg Ala Leu Glu Glu Leu Ile Ala Glu Gln
                450                 455                 460

ttg gaa gct cct atc tgc taa                                        1596
Leu Glu Ala Pro Ile Cys
            465


<210> 16
<211> 467
<212> PRT
<213> Lucilia cuprina

<400> 6
Met Asp Asn Gly Glu Gln Asp Ala Gly Phe Arg Leu Ala Pro Met Ser
 1               5                  10                  15

Pro Gln Glu Ile Lys Pro Asp Ile Ser Leu Leu Asn Glu Asn Asn Thr
             20                  25                  30
```

Ser Ser Tyr Ser Pro Lys Pro Gly Ser Pro Asn Pro Phe Ala Ile Gly
35 40 45

Leu Gln Ala Ile Asn Ala Val Ala Ala Ala Asn Ala Asn Asn Gln Asn
50 55 60

Gln Met Leu Gln Thr Thr Pro Pro Gln Gln Gln Gln Tyr Pro Pro Asn
65 70 75 80

His Pro Leu Ser Gly Ser Lys His Leu Cys Ser Ile Cys Gly Asp Arg
85 90 95

Ala Ser Gly Lys His Tyr Gly Val Tyr Ser Cys Glu Gly Cys Lys Gly
100 105 110

Phe Phe Lys Arg Thr Val Arg Lys Asp Leu Thr Tyr Ala Cys Arg Glu
115 120 125

Asp Arg Asn Cys Ile Ile Asp Lys Arg Gln Arg Asn Arg Cys Gln Tyr
130 135 140

Cys Arg Tyr Gln Lys Cys Leu Ala Cys Gly Met Lys Arg Glu Ala Val
145 150 155 160

Gln Glu Glu Arg Gln Arg Gly Thr Arg Ala Ala Asn Ala Arg Ala Ala
165 170 175

Gly Ala Gly Gly Gly Gly Gly Gly Gly Gly Gly Val Ser Asn Val Val
180 185 190

Gly Ala Gly Gly Glu Asp Phe Lys Pro Ser Ser Ser Leu Arg Asp Leu
195 200 205

Thr Ile Glu Arg Ile Ile Glu Ala Glu Gln Lys Ala Glu Ser Leu Ser
210 215 220

Gly Asp Asn Val Leu Pro Phe Leu Arg Val Gly Asn Asn Ser Met Val
225 230 235 240

Gln His Asp Tyr Lys Gly Ala Val Ser His Leu Cys Gln Met Val Asn
245 250 255

Lys Gln Leu Tyr Gln Met Val Glu Tyr Ala Arg Arg Thr Pro His Phe
260 265 270

Thr His Leu Gln Arg Glu Asp Gln Ile Leu Leu Leu Lys Ala Gly Trp

```
        275                 280                 285

Asn Glu Leu Leu Ile Ala Asn Val Ala Trp Cys Ser Ile Glu Ser Leu
    290                 295                 300

Asp Ala Glu Tyr Ala Ser Pro Gly Thr Val His Asp Gly Ser Phe Gly
305                 310                 315                 320

Arg Arg Ser Pro Val Arg Gln Pro Gln Gln Leu Phe Leu Asn Gln Asn
                325                 330                 335

Phe Ser Tyr His Arg Asn Ser Ala Ile Lys Ala Asn Val Val Ser Ile
            340                 345                 350

Phe Asp Arg Ile Leu Ser Glu Leu Ser Ile Lys Met Lys Arg Leu Asn
        355                 360                 365

Ile Asp Arg Ser Glu Leu Ser Cys Leu Lys Ala Ile Ile Leu Phe Asn
    370                 375                 380

Pro Asp Ile Arg Gly Leu Lys Cys Arg Ala Asp Val Glu Val Cys Arg
385                 390                 395                 400

Glu Lys Ile Tyr Ala Cys Leu Asp Glu His Cys Arg Thr Glu His Pro
                405                 410                 415

Gly Asp Asp Gly Arg Phe Ala Gln Leu Leu Leu Arg Leu Pro Ala Leu
            420                 425                 430

Arg Ser Ile Ser Leu Lys Cys Leu Asp His Leu Phe Phe Phe Arg Leu
        435                 440                 445

Ile Gly Glu Arg Ala Leu Glu Glu Leu Ile Ala Glu Gln Leu Glu Ala
    450                 455                 460

Pro Ile Cys
465
```

<210> 7
<211> 1536
<212> DNA
<213> Lucilia cuprina

<220>
<221> CDS

<222> (133)..(1533)

<400> 7

aattcggcac gagaaaatct tgtaataatt aaaaatcaat tatcaaaact tatagttaaa 60

tgtattaaat aaagattgtg tgtgacagaa acaaattagt gagatctctt gatacgggaa 120

aatataatca aa atg gat aac ggc gag caa gat gct ggg ttc cga ttg gca 171
Met Asp Asn Gly Glu Gln Asp Ala Gly Phe Arg Leu Ala
1 5 10

ccg atg tct ccg cag gag ata aag cca gac att tca cta ctc aat gaa 219
Pro Met Ser Pro Gln Glu Ile Lys Pro Asp Ile Ser Leu Leu Asn Glu
15 20 25

aat aat acg agt agt tat tcg ccc aaa cct gga agt cct aat cca ttt 267
Asn Asn Thr Ser Ser Tyr Ser Pro Lys Pro Gly Ser Pro Asn Pro Phe
30 35 40 45

gcc atc gga ttg cag gca ata aat gca gtc gct gcc gcg aat gcc aat 315
Ala Ile Gly Leu Gln Ala Ile Asn Ala Val Ala Ala Ala Asn Ala Asn
50 55 60

aac caa aat caa atg ttg caa act acg cca cca caa cag cag cag tat 363
Asn Gln Asn Gln Met Leu Gln Thr Thr Pro Pro Gln Gln Gln Gln Tyr
65 70 75

cca cca aat cac ccc ctt agt ggt tcg aaa cac ttg tgt tcc att tgt 411
Pro Pro Asn His Pro Leu Ser Gly Ser Lys His Leu Cys Ser Ile Cys
80 85 90

gga gac cgc gcc agt gga aaa cat tat ggg gtc tac agt tgt gag ggt 459
Gly Asp Arg Ala Ser Gly Lys His Tyr Gly Val Tyr Ser Cys Glu Gly
95 100 105

tgt aaa ggg ttc ttc aaa cgt acc gta cgc aag gac ttg aca tat gct 507
Cys Lys Gly Phe Phe Lys Arg Thr Val Arg Lys Asp Leu Thr Tyr Ala
110 115 120 125

tgt cgt gag gac aga aat tgc att ata gat aaa cga caa aga aat cgt 555
Cys Arg Glu Asp Arg Asn Cys Ile Ile Asp Lys Arg Gln Arg Asn Arg
130 135 140

tgc cag tat tgt cgt tat caa aag tgt tta gct tgt ggc atg aaa cgc 603
Cys Gln Tyr Cys Arg Tyr Gln Lys Cys Leu Ala Cys Gly Met Lys Arg
145 150 155

```
gaa gcg gtc caa gag gaa cga caa cgt ggt act cgt gct gct aac gct    651
Glu Ala Val Gln Glu Glu Arg Gln Arg Gly Thr Arg Ala Ala Asn Ala
        160                 165                 170

aga gct gct ggt gct ggc ggt ggt gga gga ggt ggt ggt ggg gta agc    699
Arg Ala Ala Gly Ala Gly Gly Gly Gly Gly Gly Gly Gly Gly Val Ser
    175                 180                 185

aat gtg gtt ggt gct ggc gga gaa gac ttt aaa ccc agc agt tca tta    747
Asn Val Val Gly Ala Gly Gly Glu Asp Phe Lys Pro Ser Ser Ser Leu
190                 195                 200                 205

cgt gat ctc act ata gaa cgc atc att gaa gcc gag caa aag gct gaa    795
Arg Asp Leu Thr Ile Glu Arg Ile Ile Glu Ala Glu Gln Lys Ala Glu
                210                 215                 220

tct ttg agc ggt gat aac gtg ttg ccc ttt ttg cgc gtt ggc aac aat    843
Ser Leu Ser Gly Asp Asn Val Leu Pro Phe Leu Arg Val Gly Asn Asn
            225                 230                 235

tcc atg gta caa cac gac tac aaa ggc gcg gta tct cat ctc tgc cag    891
Ser Met Val Gln His Asp Tyr Lys Gly Ala Val Ser His Leu Cys Gln
        240                 245                 250

atg gtt aac aaa caa ctc tac caa atg gtt gaa tat gca cgt cga aca    939
Met Val Asn Lys Gln Leu Tyr Gln Met Val Glu Tyr Ala Arg Arg Thr
    255                 260                 265

cca cat ttt aca cat ttg cag cgt gag gat cag ata cta ttg tta aag    987
Pro His Phe Thr His Leu Gln Arg Glu Asp Gln Ile Leu Leu Leu Lys
270                 275                 280                 285

gct ggc tgg aat gaa ctg cta att gca aat gtt gcc tgg tgc agt att    1035
Ala Gly Trp Asn Glu Leu Leu Ile Ala Asn Val Ala Trp Cys Ser Ile
                290                 295                 300

gag tct ctg gat gcc gaa tat gcc tct cct ggt acg gta cat gac ggt    1083
Glu Ser Leu Asp Ala Glu Tyr Ala Ser Pro Gly Thr Val His Asp Gly
            305                 310                 315

tct ttt ggt cgg cgt tca cca gtg cgt cag ccc caa caa ctc ttc ctt    1131
Ser Phe Gly Arg Arg Ser Pro Val Arg Gln Pro Gln Gln Leu Phe Leu
        320                 325                 330

aat cag aat ttc tcg tat cat cgc aat agt gct att aag gcc aat gtt    1179
```

```
Asn Gln Asn Phe Ser Tyr His Arg Asn Ser Ala Ile Lys Ala Asn Val
    335                 340                 345

gtt tca att ttc gat cgt atc ctc tcg gag ttg agc atc aaa atg aaa    1227
Val Ser Ile Phe Asp Arg Ile Leu Ser Glu Leu Ser Ile Lys Met Lys
350                 355                 360                 365

cgt ctt aac atc gat cgc tcg gag ttg tcg tgt ctg aag gca atc ata    1275
Arg Leu Asn Ile Asp Arg Ser Glu Leu Ser Cys Leu Lys Ala Ile Ile
                370                 375                 380

ctc ttc aat cca gac ata cgc ggt ctg aaa tgt cga gcc gac gtc gag    1323
Leu Phe Asn Pro Asp Ile Arg Gly Leu Lys Cys Arg Ala Asp Val Glu
            385                 390                 395

gta tgt cgt gaa aaa atc tat gcc tgt ctg gac gaa cac tgc cgc aca    1371
Val Cys Arg Glu Lys Ile Tyr Ala Cys Leu Asp Glu His Cys Arg Thr
        400                 405                 410

gaa cat cca ggt gat gat ggc cgc ttt gct cag cta cta cta agg ttg    1419
Glu His Pro Gly Asp Asp Gly Arg Phe Ala Gln Leu Leu Leu Arg Leu
    415                 420                 425

ccc gca ttg cgt tcc atc agt ctc aaa tgt ctc gat cat ttg ttt ttc    1467
Pro Ala Leu Arg Ser Ile Ser Leu Lys Cys Leu Asp His Leu Phe Phe
430                 435                 440                 445

ttc cgt tta ata ggc gaa aga gca ttg gag gaa tta att gct gag caa    1515
Phe Arg Leu Ile Gly Glu Arg Ala Leu Glu Glu Leu Ile Ala Glu Gln
                450                 455                 460

ttg gaa gct cct atc tgc taa                                        1563
Leu Glu Ala Pro Ile Cys
            465


<210> 8
<211> 467
<212> PRT
<213> Lucilia cuprina

<400> 8
Met Asp Asn Gly Glu Gln Asp Ala Gly Phe Arg Leu Ala Pro Met Ser
  1               5                  10                  15

Pro Gln Glu Ile Lys Pro Asp Ile Ser Leu Leu Asn Glu Asn Asn Thr
```

```
            20                  25                  30

Ser Ser Tyr Ser Pro Lys Pro Gly Ser Pro Asn Pro Phe Ala Ile Gly
        35                  40                  45

Leu Gln Ala Ile Asn Ala Val Ala Ala Ala Asn Ala Asn Asn Gln Asn
    50                  55                  60

Gln Met Leu Gln Thr Thr Pro Pro Gln Gln Gln Gln Tyr Pro Pro Asn
65                  70                  75                  80

His Pro Leu Ser Gly Ser Lys His Leu Cys Ser Ile Cys Gly Asp Arg
                85                  90                  95

Ala Ser Gly Lys His Tyr Gly Val Tyr Ser Cys Glu Gly Cys Lys Gly
            100                 105                 110

Phe Phe Lys Arg Thr Val Arg Lys Asp Leu Thr Tyr Ala Cys Arg Glu
        115                 120                 125

Asp Arg Asn Cys Ile Ile Asp Lys Arg Gln Arg Asn Arg Cys Gln Tyr
    130                 135                 140

Cys Arg Tyr Gln Lys Cys Leu Ala Cys Gly Met Lys Arg Glu Ala Val
145                 150                 155                 160

Gln Glu Glu Arg Gln Arg Gly Thr Arg Ala Ala Asn Ala Arg Ala Ala
                165                 170                 175

Gly Ala Gly Gly Gly Gly Gly Gly Gly Gly Gly Val Ser Asn Val Val
            180                 185                 190

Gly Ala Gly Gly Glu Asp Phe Lys Pro Ser Ser Ser Leu Arg Asp Leu
        195                 200                 205

Thr Ile Glu Arg Ile Ile Glu Ala Glu Gln Lys Ala Glu Ser Leu Ser
    210                 215                 220

Gly Asp Asn Val Leu Pro Phe Leu Arg Val Gly Asn Asn Ser Met Val
225                 230                 235                 240

Gln His Asp Tyr Lys Gly Ala Val Ser His Leu Cys Gln Met Val Asn
                245                 250                 255

Lys Gln Leu Tyr Gln Met Val Glu Tyr Ala Arg Arg Thr Pro His Phe
            260                 265                 270
```

Thr His Leu Gln Arg Glu Asp Gln Ile Leu Leu Leu Lys Ala Gly Trp
275 280 285

Asn Glu Leu Leu Ile Ala Asn Val Ala Trp Cys Ser Ile Glu Ser Leu
290 295 300

Asp Ala Glu Tyr Ala Ser Pro Gly Thr Val His Asp Gly Ser Phe Gly
305 310 315 320

Arg Arg Ser Pro Val Arg Gln Pro Gln Gln Leu Phe Leu Asn Gln Asn
325 330 335

Phe Ser Tyr His Arg Asn Ser Ala Ile Lys Ala Asn Val Val Ser Ile
340 345 350

Phe Asp Arg Ile Leu Ser Glu Leu Ser Ile Lys Met Lys Arg Leu Asn
355 360 365

Ile Asp Arg Ser Glu Leu Ser Cys Leu Lys Ala Ile Ile Leu Phe Asn
370 375 380

Pro Asp Ile Arg Gly Leu Lys Cys Arg Ala Asp Val Glu Val Cys Arg
385 390 395 400

Glu Lys Ile Tyr Ala Cys Leu Asp Glu His Cys Arg Thr Glu His Pro
405 410 415

Gly Asp Asp Gly Arg Phe Ala Gln Leu Leu Leu Arg Leu Pro Ala Leu
420 425 430

Arg Ser Ile Ser Leu Lys Cys Leu Asp His Leu Phe Phe Phe Arg Leu
435 440 445

Ile Gly Glu Arg Ala Leu Glu Glu Leu Ile Ala Glu Gln Leu Glu Ala
450 455 460

Pro Ile Cys
465

<210> 9
<211> 585
<212> DNA
<213> Myzus persicae

<220>
<221> CDS
<222> (1)..(585)

<400> 9

```
gaa ttc ggc acg agc gcc att gtt aat gga ttt atc cgc acc att agt    48
Glu Phe Gly Thr Ser Ala Ile Val Asn Gly Phe Ile Arg Thr Ile Ser
  1               5                  10                  15

ttg atc ctt att ttt ctt ctt ctt ttt ctt tgg agg ttg ttg gcc ttc    96
Leu Ile Leu Ile Phe Leu Leu Leu Phe Leu Trp Arg Leu Leu Ala Phe
             20                  25                  30

cgg ttc ttg ttt ata tct gaa caa cca cct ccc gaa gag ctg tgc ctg   144
Arg Phe Leu Phe Ile Ser Glu Gln Pro Pro Pro Glu Glu Leu Cys Leu
         35                  40                  45

gtg tgt ggc gac cgg tcg tcc ggt tac cat tac aac gct ctc aca tgc   192
Val Cys Gly Asp Arg Ser Ser Gly Tyr His Tyr Asn Ala Leu Thr Cys
     50                  55                  60

gaa gga tgc aag ggg ttc ttc cgg agg agc atc acc aag aac gcc gtg   240
Glu Gly Cys Lys Gly Phe Phe Arg Arg Ser Ile Thr Lys Asn Ala Val
 65                  70                  75                  80

tac cag tgc aag tac ggc aac aat tgc gaa atc gac atg tac atg agg   288
Tyr Gln Cys Lys Tyr Gly Asn Asn Cys Glu Ile Asp Met Tyr Met Arg
                 85                  90                  95

cgg aag tgc cag gag tgc cgg ctg aaa aaa tgc ctg acc gtc ggc atg   336
Arg Lys Cys Gln Glu Cys Arg Leu Lys Lys Cys Leu Thr Val Gly Met
            100                 105                 110

agg cct gaa tgt gtt gta cct gaa gtt caa tgc gca gta aaa aga aag   384
Arg Pro Glu Cys Val Val Pro Glu Val Gln Cys Ala Val Lys Arg Lys
        115                 120                 125

gag aaa aaa gct caa cga gaa aaa gat aaa cca aat tct act aca gac   432
Glu Lys Lys Ala Gln Arg Glu Lys Asp Lys Pro Asn Ser Thr Thr Asp
    130                 135                 140

att tct cct gaa ata ata aaa ata gaa cct aca gag atg aag att gaa   480
Ile Ser Pro Glu Ile Ile Lys Ile Glu Pro Thr Glu Met Lys Ile Glu
145                 150                 155                 160

tgt ggt gaa cca atg ata atg ggc aca cct atg ccg act gta cct tac   528
```

```
Cys Gly Glu Pro Met Ile Met Gly Thr Pro Met Pro Thr Val Pro Tyr
                165                 170                 175

gtg aaa cct ttg agt tct ctc gtg ccg aat tcg gca cga gtc acg ggt    576
Val Lys Pro Leu Ser Ser Leu Val Pro Asn Ser Ala Arg Val Thr Gly
            180                 185                 190

tac aaa ttt                                                        585
Tyr Lys Phe
        195


&lt;210&gt; 10
&lt;211&gt; 195
&lt;212&gt; PRT
&lt;213&gt; Myzus persicae

&lt;400&gt; 10
Glu Phe Gly Thr Ser Ala Ile Val Asn Gly Phe Ile Arg Thr Ile Ser
1               5                   10                  15

Leu Ile Leu Ile Phe Leu Leu Leu Phe Leu Trp Arg Leu Leu Ala Phe
            20                  25                  30

Arg Phe Leu Phe Ile Ser Glu Gln Pro Pro Pro Glu Glu Leu Cys Leu
        35                  40                  45

Val Cys Gly Asp Arg Ser Ser Gly Tyr His Tyr Asn Ala Leu Thr Cys
    50                  55                  60

Glu Gly Cys Lys Gly Phe Phe Arg Arg Ser Ile Thr Lys Asn Ala Val
65                  70                  75                  80

Tyr Gln Cys Lys Tyr Gly Asn Asn Cys Glu Ile Asp Met Tyr Met Arg
                85                  90                  95

Arg Lys Cys Gln Glu Cys Arg Leu Lys Lys Cys Leu Thr Val Gly Met
            100                 105                 110

Arg Pro Glu Cys Val Val Pro Glu Val Gln Cys Ala Val Lys Arg Lys
        115                 120                 125

Glu Lys Lys Ala Gln Arg Glu Lys Asp Lys Pro Asn Ser Thr Thr Asp
    130                 135                 140

Ile Ser Pro Glu Ile Ile Lys Ile Glu Pro Thr Glu Met Lys Ile Glu
```

145 150 155 160

Cys Gly Glu Pro Met Ile Met Gly Thr Pro Met Pro Thr Val Pro Tyr
165 170 175

Val Lys Pro Leu Ser Ser Leu Val Pro Asn Ser Ala Arg Val Thr Gly
180 185 190

Tyr Lys Phe
195

<210> 11
<211> 208
<212> DNA
<213> Myzus persicae

<400> 11
catgcctgca ggtcgactct agaggatccc ctcgtccggt taccattaca acgcactcac 60

ctgtgaaggc tgtaagggtt tctttcgacg gagtgttacc aaaaatgcgg tgtattgttg 120

taaatttggt catgcctgcg aaatggacat gtatatgcga cgtaaatgtc aggaatgtag 180

gctgaaaaaa tgtttggctg tgggcatg 208

<210> 12
<211> 436
<212> DNA
<213> Myzus persicae

<400> 12
catgcggccg gaatgtgtgg tgcccgaaaa ccagtgtgca atgaaacgac gcgaaaagaa 60

agcacaaaaa gagaaggata aaatacagac cagtgtgtgt gcaacggaaa ttaaaaagga 120

aatactcgat ttaatgacat gtgaaccgcc atcacatcca acgtgtccgc tgttacctga 180

agacattttg gctaaatgtc aagctcgtaa tatacctcct ttatcgtaca atcaattggc 240

agttatatat aaattaatat ggtatcaaga tggctacgaa cagccatccg aggaagatct 300

caaacgtata atgagttcac ccgatgaaaa tgaaagtcaa cacgatgcat catttcgtca 360

tataacagaa atcactatac taacagtaca attaattgtt gaatgtgcca aaggtctagg 420

```
gtaccgagct cgaatt                                                  436




<210> 13
<211> 1797
<212> DNA
<213> Myzus persicae

<220>
<221> CDS
<222> (1)..(1797)

<400> 13
atg atg gac cag aaa tgt gac gtc ggc ggt ggt ggt gtc gct gct gcc     48
Met Met Asp Gln Lys Cys Asp Val Gly Gly Gly Gly Val Ala Ala Ala
  1               5                  10                  15

gcc gcc ggt atc ggt ggc ggc ggt gtc ggc ggc ctc atg tcg tac aac     96
Ala Ala Gly Ile Gly Gly Gly Gly Val Gly Gly Leu Met Ser Tyr Asn
             20                  25                  30

cgt ggc cgt ggc ggc acc gag gtc atc atc aaa ccc cgt agt cct gcc    144
Arg Gly Arg Gly Gly Thr Glu Val Ile Ile Lys Pro Arg Ser Pro Ala
         35                  40                  45

gtg gtg cag gtg gcc acc ggt ggc agt tac cac ggc ctg ccg gcg gcc    192
Val Val Gln Val Ala Thr Gly Gly Ser Tyr His Gly Leu Pro Ala Ala
     50                  55                  60

tcc gac gcc gtc atc gtg cgc agc ccg cca ggc ggc cac ttg ccc ggg    240
Ser Asp Ala Val Ile Val Arg Ser Pro Pro Gly Gly His Leu Pro Gly
 65                  70                  75                  80

ccg cag cag caa gtg ccg ccg tcc cgc aac ggc tgt tcc acc ctg ttt    288
Pro Gln Gln Gln Val Pro Pro Ser Arg Asn Gly Cys Ser Thr Leu Phe
                 85                  90                  95

agc gac atc gct ggc gtc aag cga ctc agg ccc gac gat tgg ttg gcc    336
Ser Asp Ile Ala Gly Val Lys Arg Leu Arg Pro Asp Asp Trp Leu Ala
            100                 105                 110

gtc aac tcg ccg ccc gcc tct tcg ccc ggc acg tcg cac ata tcc tac    384
Val Asn Ser Pro Pro Ala Ser Ser Pro Gly Thr Ser His Ile Ser Tyr
        115                 120                 125
```

```
aca gtc ata tcg aac ggc ggc ggc ggt ggc ggc ggt ggc ggc ggt ggt    432
Thr Val Ile Ser Asn Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly
    130                 135                 140

tac aac acg tct cca atg tcg acc aac agc tac gac ccg tac agt ccg    480
Tyr Asn Thr Ser Pro Met Ser Thr Asn Ser Tyr Asp Pro Tyr Ser Pro
145                 150                 155                 160

atg agt gga aaa atc gtc aaa gaa gag ttg tct ccg cca aac agc ctg    528
Met Ser Gly Lys Ile Val Lys Glu Glu Leu Ser Pro Pro Asn Ser Leu
                165                 170                 175

tcg gga gtc agc agc cat tcg gat ggg ttg aag aag aag aaa ctc aac    576
Ser Gly Val Ser Ser His Ser Asp Gly Leu Lys Lys Lys Lys Leu Asn
            180                 185                 190

cac acg ccc tcg acc ggt gtc gtc aac acc tcg gca tcg ggc ccc ggg    624
His Thr Pro Ser Thr Gly Val Val Asn Thr Ser Ala Ser Gly Pro Gly
        195                 200                 205

ggt ggc gtt ggt ggc aat gtg ctg aac aac cga cct ccc gaa gag ctg    672
Gly Gly Val Gly Gly Asn Val Leu Asn Asn Arg Pro Pro Glu Glu Leu
    210                 215                 220

tgc ctg gtg tgt ggc gac cgg tcg tcc ggt tac cat tac aac gct ctc    720
Cys Leu Val Cys Gly Asp Arg Ser Ser Gly Tyr His Tyr Asn Ala Leu
225                 230                 235                 240

aca tgc gaa gga tgc aag ggg ttc ttc cgg agg agc atc acc aag aac    768
Thr Cys Glu Gly Cys Lys Gly Phe Phe Arg Arg Ser Ile Thr Lys Asn
                245                 250                 255

gcc gtg tac cag tgc aag tac ggc aac aat tgc gaa atc gac atg tac    816
Ala Val Tyr Gln Cys Lys Tyr Gly Asn Asn Cys Glu Ile Asp Met Tyr
            260                 265                 270

atg agg cgg aag tgc cag gag tgc cgg ctg aaa aaa tgc ctg acc gtc    864
Met Arg Arg Lys Cys Gln Glu Cys Arg Leu Lys Lys Cys Leu Thr Val
        275                 280                 285

ggc atg agg cct gaa tgt gtt gta cct gaa gtt caa tgc gca gta aaa    912
Gly Met Arg Pro Glu Cys Val Val Pro Glu Val Gln Cys Ala Val Lys
    290                 295                 300

aga aag gag aaa aaa gct caa cga gaa aaa gat aaa cca aat tct act    960
```

```
Arg Lys Glu Lys Lys Ala Gln Arg Glu Lys Asp Lys Pro Asn Ser Thr
305                 310                 315                 320

aca gac att tct cct gaa ata ata aaa ata gaa cct aca gag atg aag     1008
Thr Asp Ile Ser Pro Glu Ile Ile Lys Ile Glu Pro Thr Glu Met Lys
                325                 330                 335

att gaa tgt ggt gaa cca atg ata atg ggc aca cct atg ccg act gta     1056
Ile Glu Cys Gly Glu Pro Met Ile Met Gly Thr Pro Met Pro Thr Val
            340                 345                 350

cct tac gtg aaa cct ttg agt tct gaa caa aaa gaa ctg atc cac cga     1104
Pro Tyr Val Lys Pro Leu Ser Ser Glu Gln Lys Glu Leu Ile His Arg
        355                 360                 365

ctt gtc tat ttc cag gat caa tat gaa gct cct agt gaa aaa gac atg     1152
Leu Val Tyr Phe Gln Asp Gln Tyr Glu Ala Pro Ser Glu Lys Asp Met
    370                 375                 380

aaa cgt tta aca ata aat aat caa aat atg gat gaa tat gat gaa gaa     1200
Lys Arg Leu Thr Ile Asn Asn Gln Asn Met Asp Glu Tyr Asp Glu Glu
385                 390                 395                 400

aaa caa agt gac acc aca tat cga atc atc act gag atg aca ata ctc     1248
Lys Gln Ser Asp Thr Thr Tyr Arg Ile Ile Thr Glu Met Thr Ile Leu
                405                 410                 415

aca gtt caa ctg att gtt gag ttt gcc aaa cga tta cca ggt ttc gat     1296
Thr Val Gln Leu Ile Val Glu Phe Ala Lys Arg Leu Pro Gly Phe Asp
            420                 425                 430

aaa ctt gta aga gaa gat caa atc act tta ctc aag gct tgc tca agt     1344
Lys Leu Val Arg Glu Asp Gln Ile Thr Leu Leu Lys Ala Cys Ser Ser
        435                 440                 445

gaa gct atg atg ttc agg gta gca agg aag tat gac atc acc act gac     1392
Glu Ala Met Met Phe Arg Val Ala Arg Lys Tyr Asp Ile Thr Thr Asp
    450                 455                 460

tca ata gtg ttt gct aac aac cag cca ttt tca gct gat tca tat aac     1440
Ser Ile Val Phe Ala Asn Asn Gln Pro Phe Ser Ala Asp Ser Tyr Asn
465                 470                 475                 480

aaa gct gga ttg gga gat gcc att gaa aac caa ctg tca ttc agt cgg     1488
Lys Ala Gly Leu Gly Asp Ala Ile Glu Asn Gln Leu Ser Phe Ser Arg
                485                 490                 495
```

```
ttt atg tac aat atg aag gtg gat aac gca gaa tat gcc tta ttg acc    1536
Phe Met Tyr Asn Met Lys Val Asp Asn Ala Glu Tyr Ala Leu Leu Thr
            500                 505                 510

gcc atc gtc ata ttt tcg agt agg cca aat tta cta gat ggt tgg aaa    1584
Ala Ile Val Ile Phe Ser Ser Arg Pro Asn Leu Leu Asp Gly Trp Lys
        515                 520                 525

gtg gag aaa atc caa gaa atc tac cta gag tcc tta aaa gct tat gta    1632
Val Glu Lys Ile Gln Glu Ile Tyr Leu Glu Ser Leu Lys Ala Tyr Val
    530                 535                 540

gat aat cga gac cgt gac aca gca act gta cga tat gcg cga ctt ctc    1680
Asp Asn Arg Asp Arg Asp Thr Ala Thr Val Arg Tyr Ala Arg Leu Leu
545                 550                 555                 560

tca gta ctt aca gaa ttg cgc aca tta ggc aat gaa aac tct gag cta    1728
Ser Val Leu Thr Glu Leu Arg Thr Leu Gly Asn Glu Asn Ser Glu Leu
                565                 570                 575

tgt atg aca ctg aaa ctg aaa aac aga gta gta ccc cca ttc ttg gcc    1776
Cys Met Thr Leu Lys Leu Lys Asn Arg Val Val Pro Pro Phe Leu Ala
            580                 585                 590

gaa ata tgg gat gtc atg cca                                        1797
Glu Ile Trp Asp Val Met Pro
        595


<210> 14
<211> 599
<212> PRT
<213> Myzus persicae


<400> 14
Met Met Asp Gln Lys Cys Asp Val Gly Gly Gly Gly Val Ala Ala Ala
1               5                   10                  15

Ala Ala Gly Ile Gly Gly Gly Gly Val Gly Gly Leu Met Ser Tyr Asn
            20                  25                  30

Arg Gly Arg Gly Gly Thr Glu Val Ile Ile Lys Pro Arg Ser Pro Ala
        35                  40                  45

Val Val Gln Val Ala Thr Gly Gly Ser Tyr His Gly Leu Pro Ala Ala
```

```
    50                  55                  60

Ser Asp Ala Val Ile Val Arg Ser Pro Pro Gly Gly His Leu Pro Gly
 65                 70                  75                  80

Pro Gln Gln Gln Val Pro Pro Ser Arg Asn Gly Cys Ser Thr Leu Phe
                85                  90                  95

Ser Asp Ile Ala Gly Val Lys Arg Leu Arg Pro Asp Asp Trp Leu Ala
            100                 105                 110

Val Asn Ser Pro Pro Ala Ser Ser Pro Gly Thr Ser His Ile Ser Tyr
        115                 120                 125

Thr Val Ile Ser Asn Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly
    130                 135                 140

Tyr Asn Thr Ser Pro Met Ser Thr Asn Ser Tyr Asp Pro Tyr Ser Pro
145                 150                 155                 160

Met Ser Gly Lys Ile Val Lys Glu Glu Leu Ser Pro Pro Asn Ser Leu
                165                 170                 175

Ser Gly Val Ser Ser His Ser Asp Gly Leu Lys Lys Lys Lys Leu Asn
            180                 185                 190

His Thr Pro Ser Thr Gly Val Val Asn Thr Ser Ala Ser Gly Pro Gly
        195                 200                 205

Gly Gly Val Gly Gly Asn Val Leu Asn Asn Arg Pro Pro Glu Glu Leu
    210                 215                 220

Cys Leu Val Cys Gly Asp Arg Ser Ser Gly Tyr His Tyr Asn Ala Leu
225                 230                 235                 240

Thr Cys Glu Gly Cys Lys Gly Phe Phe Arg Arg Ser Ile Thr Lys Asn
                245                 250                 255

Ala Val Tyr Gln Cys Lys Tyr Gly Asn Asn Cys Glu Ile Asp Met Tyr
            260                 265                 270

Met Arg Arg Lys Cys Gln Glu Cys Arg Leu Lys Lys Cys Leu Thr Val
        275                 280                 285

Gly Met Arg Pro Glu Cys Val Val Pro Glu Val Gln Cys Ala Val Lys
    290                 295                 300
```

Arg Lys Glu Lys Lys Ala Gln Arg Glu Lys Asp Lys Pro Asn Ser Thr
305 310 315 320

Thr Asp Ile Ser Pro Glu Ile Ile Lys Ile Glu Pro Thr Glu Met Lys
325 330 335

Ile Glu Cys Gly Glu Pro Met Ile Met Gly Thr Pro Met Pro Thr Val
340 345 350

Pro Tyr Val Lys Pro Leu Ser Ser Glu Gln Lys Glu Leu Ile His Arg
355 360 365

Leu Val Tyr Phe Gln Asp Gln Tyr Glu Ala Pro Ser Glu Lys Asp Met
370 375 380

Lys Arg Leu Thr Ile Asn Asn Gln Asn Met Asp Glu Tyr Asp Glu Glu
385 390 395 400

Lys Gln Ser Asp Thr Thr Tyr Arg Ile Ile Thr Glu Met Thr Ile Leu
405 410 415

Thr Val Gln Leu Ile Val Glu Phe Ala Lys Arg Leu Pro Gly Phe Asp
420 425 430

Lys Leu Val Arg Glu Asp Gln Ile Thr Leu Leu Lys Ala Cys Ser Ser
435 440 445

Glu Ala Met Met Phe Arg Val Ala Arg Lys Tyr Asp Ile Thr Thr Asp
450 455 460

Ser Ile Val Phe Ala Asn Asn Gln Pro Phe Ser Ala Asp Ser Tyr Asn
465 470 475 480

Lys Ala Gly Leu Gly Asp Ala Ile Glu Asn Gln Leu Ser Phe Ser Arg
485 490 495

Phe Met Tyr Asn Met Lys Val Asp Asn Ala Glu Tyr Ala Leu Leu Thr
500 505 510

Ala Ile Val Ile Phe Ser Ser Arg Pro Asn Leu Leu Asp Gly Trp Lys
515 520 525

Val Glu Lys Ile Gln Glu Ile Tyr Leu Glu Ser Leu Lys Ala Tyr Val
530 535 540

Asp Asn Arg Asp Arg Asp Thr Ala Thr Val Arg Tyr Ala Arg Leu Leu
545 550 555 560

Ser Val Leu Thr Glu Leu Arg Thr Leu Gly Asn Glu Asn Ser Glu Leu
565 570 575

Cys Met Thr Leu Lys Leu Lys Asn Arg Val Val Pro Pro Phe Leu Ala
580 585 590

Glu Ile Trp Asp Val Met Pro
595

<210> 15
<211> 1131
<212> DNA
<213> Myzus persicae
<220>
<221> CDS
<222> (1)..(1131)

<400> 15
atg tat tcc aac tcg tac acc atg tat tca agt gac aga tta tac agc 48
Met Tyr Ser Asn Ser Tyr Thr Met Tyr Ser Ser Asp Arg Leu Tyr Ser
1 5 10 15

gtc gat cgg aac agt atg atg aat aat tct tgc aac gta caa gac tct 96
Val Asp Arg Asn Ser Met Met Asn Asn Ser Cys Asn Val Gln Asp Ser
20 25 30

ccg aat tac ccg ccc aac cat cca ctc agc ggt tcg aaa cat ctg tgc 144
Pro Asn Tyr Pro Pro Asn His Pro Leu Ser Gly Ser Lys His Leu Cys
35 40 45

tcc ata tgc ggc gat cgc gcc agt gga aaa cat tac gga gtc tac agc 192
Ser Ile Cys Gly Asp Arg Ala Ser Gly Lys His Tyr Gly Val Tyr Ser
50 55 60

tgc gag ggg tgc aaa ggg ttc ttc aaa cgc aca gtg agg aaa aat ttg 240
Cys Glu Gly Cys Lys Gly Phe Phe Lys Arg Thr Val Arg Lys Asn Leu
65 70 75 80

tca tac gcg tgt cgc gaa gaa aac aaa tgc atc atc gac aag cgc caa 288
Ser Tyr Ala Cys Arg Glu Glu Asn Lys Cys Ile Ile Asp Lys Arg Gln
85 90 95

```
cga aat cgg tgc caa tac tgc agg tat caa aaa tgt ttg acc atg ggc    336
Arg Asn Arg Cys Gln Tyr Cys Arg Tyr Gln Lys Cys Leu Thr Met Gly
            100                 105                 110

atg aaa aga gaa gct gtg cag gaa gaa aga caa cgt aca aaa gaa cga    384
Met Lys Arg Glu Ala Val Gln Glu Glu Arg Gln Arg Thr Lys Glu Arg
        115                 120                 125

gat cat aat aac atc gaa gtt gaa ccc acg agc agt tct aat act gat    432
Asp His Asn Asn Ile Glu Val Glu Pro Thr Ser Ser Ser Asn Thr Asp
    130                 135                 140

atg cca gtg gaa ctc ata tta agg gct gag aat aaa gct gat gct ata    480
Met Pro Val Glu Leu Ile Leu Arg Ala Glu Asn Lys Ala Asp Ala Ile
145                 150                 155                 160

aag act gaa caa cag tat ata gag caa cga cat cct caa cat act gtt    528
Lys Thr Glu Gln Gln Tyr Ile Glu Gln Arg His Pro Gln His Thr Val
                165                 170                 175

ggt gct att tgt caa gca act gac aag cag tta ata caa ctt gtt gaa    576
Gly Ala Ile Cys Gln Ala Thr Asp Lys Gln Leu Ile Gln Leu Val Glu
            180                 185                 190

tgg gcc aag cat ata ccg cat ttt aaa aat tta cct cta ggc gat caa    624
Trp Ala Lys His Ile Pro His Phe Lys Asn Leu Pro Leu Gly Asp Gln
        195                 200                 205

gtt tta tta ttg aga gct ggt tgg aat gag ttg atg att gca gca ttt    672
Val Leu Leu Leu Arg Ala Gly Trp Asn Glu Leu Met Ile Ala Ala Phe
    210                 215                 220

tcc cat aga tca atc agt gta aaa gat ggt ata gtc tta gct act gga    720
Ser His Arg Ser Ile Ser Val Lys Asp Gly Ile Val Leu Ala Thr Gly
225                 230                 235                 240

ctt act gtt gac aga gat tca gct cac caa gct ggt gtt gaa gct ata    768
Leu Thr Val Asp Arg Asp Ser Ala His Gln Ala Gly Val Glu Ala Ile
                245                 250                 255

ttt gat cgt gta ctc act gaa ctc gtt gct aaa atg aga gat atg ggt    816
Phe Asp Arg Val Leu Thr Glu Leu Val Ala Lys Met Arg Asp Met Gly
            260                 265                 270

atg gat aga aca gag ctt ggc tgt ttg cgt act att att ctt ttt aat    864
Met Asp Arg Thr Glu Leu Gly Cys Leu Arg Thr Ile Ile Leu Phe Asn
```

275 280 285

cca ggt tca aaa ggt ttg cag tct gtg aat gaa gtg caa gta ctg cgt 912
Pro Gly Ser Lys Gly Leu Gln Ser Val Asn Glu Val Gln Val Leu Arg
290 295 300

gat aag gtt tat gtt gcg tta gaa gaa tat tgt cgt aca aca cat cca 960
Asp Lys Val Tyr Val Ala Leu Glu Glu Tyr Cys Arg Thr Thr His Pro
305 310 315 320

gaa gaa cct gga cga ttt gct aaa cta ctt ctt cgg ctt cct tca tta 1008
Glu Glu Pro Gly Arg Phe Ala Lys Leu Leu Leu Arg Leu Pro Ser Leu
325 330 335

cgt tca att gga tta aaa tgt ctg gaa cat tta ttc ttt tat aaa ctt 1056
Arg Ser Ile Gly Leu Lys Cys Leu Glu His Leu Phe Phe Tyr Lys Leu
340 345 350

att ggc gat tcc cca att gat aca ttt tta atg gaa gtt ctc gaa tca 1104
Ile Gly Asp Ser Pro Ile Asp Thr Phe Leu Met Glu Val Leu Glu Ser
355 360 365

tct tca cat gac gtt caa gta gct aca 1131
Ser Ser His Asp Val Gln Val Ala Thr
370 375

<210> 16
<211> 377
<212> PRT
<213> Myzus persicae

<400> 16
Met Tyr Ser Asn Ser Tyr Thr Met Tyr Ser Ser Asp Arg Leu Tyr Ser
1 5 10 15

Val Asp Arg Asn Ser Met Met Asn Asn Ser Cys Asn Val Gln Asp Ser
20 25 30

Pro Asn Tyr Pro Pro Asn His Pro Leu Ser Gly Ser Lys His Leu Cys
35 40 45

Ser Ile Cys Gly Asp Arg Ala Ser Gly Lys His Tyr Gly Val Tyr Ser
50 55 60

Cys Glu Gly Cys Lys Gly Phe Phe Lys Arg Thr Val Arg Lys Asn Leu

65 70 75 80

Ser Tyr Ala Cys Arg Glu Glu Asn Lys Cys Ile Ile Asp Lys Arg Gln
85 90 95

Arg Asn Arg Cys Gln Tyr Cys Arg Tyr Gln Lys Cys Leu Thr Met Gly
100 105 110

Met Lys Arg Glu Ala Val Gln Glu Glu Arg Gln Arg Thr Lys Glu Arg
115 120 125

Asp His Asn Asn Ile Glu Val Glu Pro Thr Ser Ser Ser Asn Thr Asp
130 135 140

Met Pro Val Glu Leu Ile Leu Arg Ala Glu Asn Lys Ala Asp Ala Ile
145 150 155 160

Lys Thr Glu Gln Gln Tyr Ile Glu Gln Arg His Pro Gln His Thr Val
165 170 175

Gly Ala Ile Cys Gln Ala Thr Asp Lys Gln Leu Ile Gln Leu Val Glu
180 185 190

Trp Ala Lys His Ile Pro His Phe Lys Asn Leu Pro Leu Gly Asp Gln
195 200 205

Val Leu Leu Leu Arg Ala Gly Trp Asn Glu Leu Met Ile Ala Ala Phe
210 215 220

Ser His Arg Ser Ile Ser Val Lys Asp Gly Ile Val Leu Ala Thr Gly
225 230 235 240

Leu Thr Val Asp Arg Asp Ser Ala His Gln Ala Gly Val Glu Ala Ile
245 250 255

Phe Asp Arg Val Leu Thr Glu Leu Val Ala Lys Met Arg Asp Met Gly
260 265 270

Met Asp Arg Thr Glu Leu Gly Cys Leu Arg Thr Ile Ile Leu Phe Asn
275 280 285

Pro Gly Ser Lys Gly Leu Gln Ser Val Asn Glu Val Gln Val Leu Arg
290 295 300

Asp Lys Val Tyr Val Ala Leu Glu Glu Tyr Cys Arg Thr Thr His Pro
305 310 315 320

Glu Glu Pro Gly Arg Phe Ala Lys Leu Leu Leu Arg Leu Pro Ser Leu
325 330 335

Arg Ser Ile Gly Leu Lys Cys Leu Glu His Leu Phe Phe Tyr Lys Leu
340 345 350

Ile Gly Asp Ser Pro Ile Asp Thr Phe Leu Met Glu Val Leu Glu Ser
355 360 365

Ser Ser His Asp Val Gln Val Ala Thr
370 375

<210> 17
<211> 1131
<212> DNA
<213> Myzus persicae

<220>
<221> CDS
<222> (1)..(1131)

<400> 17
atg tat tcc aac tcg tac acc atg tat tca agt gac aga tta tac agc 48
Met Tyr Ser Asn Ser Tyr Thr Met Tyr Ser Ser Asp Arg Leu Tyr Ser
1 5 10 15

gtc gat cgg aac agt atg atg aat aat tct tgc aac gta caa gac tct 96
Val Asp Arg Asn Ser Met Met Asn Asn Ser Cys Asn Val Gln Asp Ser
20 25 30

ccg aat tac ccg ccc aac cat cca ctc agc ggt tcg aaa cat ctg tgc 144
Pro Asn Tyr Pro Pro Asn His Pro Leu Ser Gly Ser Lys His Leu Cys
35 40 45

tcc ata tgc ggc gat cgc gcc agt gga aaa cat tac gga gtc tac agc 192
Ser Ile Cys Gly Asp Arg Ala Ser Gly Lys His Tyr Gly Val Tyr Ser
50 55 60

tgc gag ggg tgc aaa ggg ttc ttc aaa cgc aca gtg agg aaa aat ttg 240
Cys Glu Gly Cys Lys Gly Phe Phe Lys Arg Thr Val Arg Lys Asn Leu
65 70 75 80

tca tac gcg tgt cgc gaa gaa aac aaa tgc atc atc gac aag cgc caa 288
Ser Tyr Ala Cys Arg Glu Glu Asn Lys Cys Ile Ile Asp Lys Arg Gln
85 90 95

```
cga aat cgg tgc caa tac tgc agg tat caa aaa tgt ttg acc atg ggc    336
Arg Asn Arg Cys Gln Tyr Cys Arg Tyr Gln Lys Cys Leu Thr Met Gly
            100                 105                 110

atg aaa aga gaa gct gtg cag gaa gaa aga caa cgt aca aaa gaa cga    384
Met Lys Arg Glu Ala Val Gln Glu Glu Arg Gln Arg Thr Lys Glu Arg
        115                 120                 125

gat cat aat aac atc gaa gtt gaa ccc acg agc agt tct aat act gat    432
Asp His Asn Asn Ile Glu Val Glu Pro Thr Ser Ser Ser Asn Thr Asp
    130                 135                 140

atg cca gtg gaa ctc ata tta agg gct gag aat aaa gct gat gct ata    480
Met Pro Val Glu Leu Ile Leu Arg Ala Glu Asn Lys Ala Asp Ala Ile
145                 150                 155                 160

aag act gaa caa cag tat ata gag caa cga cat cct caa cat act gtt    528
Lys Thr Glu Gln Gln Tyr Ile Glu Gln Arg His Pro Gln His Thr Val
                165                 170                 175

ggt gct att tgt caa gca act gac aag cag tta ata caa ctt gtt gaa    576
Gly Ala Ile Cys Gln Ala Thr Asp Lys Gln Leu Ile Gln Leu Val Glu
            180                 185                 190

tgg gcc aag cat ata ccg cat ttt aaa aat tta cct cta ggc gat caa    624
Trp Ala Lys His Ile Pro His Phe Lys Asn Leu Pro Leu Gly Asp Gln
        195                 200                 205

gtt tta tta ttg aga gct ggt tgg aat gag ttg atg att gca gca ttt    672
Val Leu Leu Leu Arg Ala Gly Trp Asn Glu Leu Met Ile Ala Ala Phe
    210                 215                 220

tcc cat aga tca atc agt gta aaa gat ggt ata gtc tta gct act gga    720
Ser His Arg Ser Ile Ser Val Lys Asp Gly Ile Val Leu Ala Thr Gly
225                 230                 235                 240

ctt act gtt gac aga gat tca gct cac caa gct ggt gtt gaa gct ata    768
Leu Thr Val Asp Arg Asp Ser Ala His Gln Ala Gly Val Glu Ala Ile
                245                 250                 255

ttt gat cgt gta ctc act gaa ctc gtt gct aaa atg aga gat atg ggt    816
Phe Asp Arg Val Leu Thr Glu Leu Val Ala Lys Met Arg Asp Met Gly
            260                 265                 270

atg gat aga aca gag ctt ggc tgt ttg cgt act att att ctt ttt aat    864
Met Asp Arg Thr Glu Leu Gly Cys Leu Arg Thr Ile Ile Leu Phe Asn
```

275 280 285

cca ggt tca aaa ggt ttg cag tct gtg aat gaa gtg gaa gta ctg cgt 912
Pro Gly Ser Lys Gly Leu Gln Ser Val Asn Glu Val Glu Val Leu Arg
290 295 300

gat aag gtt tat gtt gcg tta gaa gaa tat tgt cgt aca aca cat cca 960
Asp Lys Val Tyr Val Ala Leu Glu Glu Tyr Cys Arg Thr Thr His Pro
305 310 315 320

gaa gaa cct gga cga ttt gct aaa cta ctt ctt cgg ctt cct tca tta 1008
Glu Glu Pro Gly Arg Phe Ala Lys Leu Leu Leu Arg Leu Pro Ser Leu
325 330 335

cgt tca att gga tta aaa tgt ctg gaa cat tta ttc ttt tat aaa ctt 1056
Arg Ser Ile Gly Leu Lys Cys Leu Glu His Leu Phe Phe Tyr Lys Leu
340 345 350

att ggc gat tcc cca att gat aca ttt tta atg gaa gtt ctc gaa tca 1104
Ile Gly Asp Ser Pro Ile Asp Thr Phe Leu Met Glu Val Leu Glu Ser
355 360 365

tct tca cat gac gtt caa gta gct aca 1131
Ser Ser His Asp Val Gln Val Ala Thr
370 375

<210> 18
<211> 377
<212> PRT
<213> Myzus persicae

<400> 18
Met Tyr Ser Asn Ser Tyr Thr Met Tyr Ser Ser Asp Arg Leu Tyr Ser
1 5 10 15

Val Asp Arg Asn Ser Met Met Asn Asn Ser Cys Asn Val Gln Asp Ser
20 25 30

Pro Asn Tyr Pro Pro Asn His Pro Leu Ser Gly Ser Lys His Leu Cys
35 40 45

Ser Ile Cys Gly Asp Arg Ala Ser Gly Lys His Tyr Gly Val Tyr Ser
50 55 60

Cys Glu Gly Cys Lys Gly Phe Phe Lys Arg Thr Val Arg Lys Asn Leu

```
65                  70                  75                  80

Ser Tyr Ala Cys Arg Glu Glu Asn Lys Cys Ile Ile Asp Lys Arg Gln
                85                  90                  95

Arg Asn Arg Cys Gln Tyr Cys Arg Tyr Gln Lys Cys Leu Thr Met Gly
            100                 105                 110

Met Lys Arg Glu Ala Val Gln Glu Glu Arg Gln Arg Thr Lys Glu Arg
        115                 120                 125

Asp His Asn Asn Ile Glu Val Glu Pro Thr Ser Ser Ser Asn Thr Asp
    130                 135                 140

Met Pro Val Glu Leu Ile Leu Arg Ala Glu Asn Lys Ala Asp Ala Ile
145                 150                 155                 160

Lys Thr Glu Gln Gln Tyr Ile Glu Gln Arg His Pro Gln His Thr Val
                165                 170                 175

Gly Ala Ile Cys Gln Ala Thr Asp Lys Gln Leu Ile Gln Leu Val Glu
            180                 185                 190

Trp Ala Lys His Ile Pro His Phe Lys Asn Leu Pro Leu Gly Asp Gln
        195                 200                 205

Val Leu Leu Leu Arg Ala Gly Trp Asn Glu Leu Met Ile Ala Ala Phe
    210                 215                 220

Ser His Arg Ser Ile Ser Val Lys Asp Gly Ile Val Leu Ala Thr Gly
225                 230                 235                 240

Leu Thr Val Asp Arg Asp Ser Ala His Gln Ala Gly Val Glu Ala Ile
                245                 250                 255

Phe Asp Arg Val Leu Thr Glu Leu Val Ala Lys Met Arg Asp Met Gly
            260                 265                 270

Met Asp Arg Thr Glu Leu Gly Cys Leu Arg Thr Ile Ile Leu Phe Asn
        275                 280                 285

Pro Gly Ser Lys Gly Leu Gln Ser Val Asn Glu Val Glu Val Leu Arg
    290                 295                 300

Asp Lys Val Tyr Val Ala Leu Glu Glu Tyr Cys Arg Thr Thr His Pro
305                 310                 315                 320
```

Glu Glu Pro Gly Arg Phe Ala Lys Leu Leu Leu Arg Leu Pro Ser Leu
325 330 335

Arg Ser Ile Gly Leu Lys Cys Leu Glu His Leu Phe Phe Tyr Lys Leu
340 345 350

Ile Gly Asp Ser Pro Ile Asp Thr Phe Leu Met Glu Val Leu Glu Ser
355 360 365

Ser Ser His Asp Val Gln Val Ala Thr
370 375

<210> 19
<211> 1242
<212> DNA
<213> Myzus persicae

<220>
<221> CDS
<222> (1)..(1239)

<400> 19
atg gac ggc acc gaa cga gga tta aga ttg gac aat aat ctg tct ctg 48
Met Asp Gly Thr Glu Arg Gly Leu Arg Leu Asp Asn Asn Leu Ser Leu
1 5 10 15

agt tca atg ggt cct cag tcg ccc cta gac ctc aaa cct gac acg gca 96
Ser Ser Met Gly Pro Gln Ser Pro Leu Asp Leu Lys Pro Asp Thr Ala
20 25 30

act tta atg gtt aat ttc agt cct ccg gga gct cct cta agt cct gca 144
Thr Leu Met Val Asn Phe Ser Pro Pro Gly Ala Pro Leu Ser Pro Ala
35 40 45

gga tta tac agc gtc gat cgg aac agt atg atg aat aat tct tgc aac 192
Gly Leu Tyr Ser Val Asp Arg Asn Ser Met Met Asn Asn Ser Cys Asn
50 55 60

gta caa gac tct ccg aat tac ccg ccc aac cat cca ctc agc ggt tcg 240
Val Gln Asp Ser Pro Asn Tyr Pro Pro Asn His Pro Leu Ser Gly Ser
65 70 75 80

aaa cat ctg tgc tcc ata tgc ggc gat cgc gcc agt gga aaa cat tac 288
Lys His Leu Cys Ser Ile Cys Gly Asp Arg Ala Ser Gly Lys His Tyr

```
                85                  90                  95

gga gtc tac agc tgc gag ggg tgc aaa ggg ttc ttc aaa cgc aca gtg     336
Gly Val Tyr Ser Cys Glu Gly Cys Lys Gly Phe Phe Lys Arg Thr Val
            100                 105                 110

agg aaa aat ttg tca tac gcg tgt cgc gaa gaa aac aaa tgc atc atc     384
Arg Lys Asn Leu Ser Tyr Ala Cys Arg Glu Glu Asn Lys Cys Ile Ile
        115                 120                 125

gac aag cgc caa cga aat cgg tgc caa tac tgc agg tat caa aaa tgt     432
Asp Lys Arg Gln Arg Asn Arg Cys Gln Tyr Cys Arg Tyr Gln Lys Cys
    130                 135                 140

ttg acc atg ggc atg aaa aga gaa gct gtg cag gaa gaa aga caa cgt     480
Leu Thr Met Gly Met Lys Arg Glu Ala Val Gln Glu Glu Arg Gln Arg
145                 150                 155                 160

aca aaa gaa cga gat cat aat aac atc gaa gtt gaa ccc acg agc agt     528
Thr Lys Glu Arg Asp His Asn Asn Ile Glu Val Glu Pro Thr Ser Ser
                165                 170                 175

tct aat act gat atg cca gtg gaa ctc ata tta agg gct gag aat aaa     576
Ser Asn Thr Asp Met Pro Val Glu Leu Ile Leu Arg Ala Glu Asn Lys
            180                 185                 190

gct gat gct ata aag act gaa caa cag tat ata gag caa cga cat cct     624
Ala Asp Ala Ile Lys Thr Glu Gln Gln Tyr Ile Glu Gln Arg His Pro
        195                 200                 205

caa cat act gtt ggt gct att tgt caa gca act gac aag cag tta ata     672
Gln His Thr Val Gly Ala Ile Cys Gln Ala Thr Asp Lys Gln Leu Ile
    210                 215                 220

caa ctt gtt gaa tgg gcc aag cat ata ccg cat ttt aaa aat tta cct     720
Gln Leu Val Glu Trp Ala Lys His Ile Pro His Phe Lys Asn Leu Pro
225                 230                 235                 240

cta ggc gat caa gtt tta tta ttg aga gct ggt tgg aat gag ttg atg     768
Leu Gly Asp Gln Val Leu Leu Leu Arg Ala Gly Trp Asn Glu Leu Met
                245                 250                 255

att gca gca ttt tcc cat aga tca atc agt gta aaa gat ggt ata gtc     816
Ile Ala Ala Phe Ser His Arg Ser Ile Ser Val Lys Asp Gly Ile Val
            260                 265                 270
```

tta gct act gga ctt act gtt gac aga gat tca gct cac caa gct ggt 864
Leu Ala Thr Gly Leu Thr Val Asp Arg Asp Ser Ala His Gln Ala Gly
275 280 285

gtt gaa gct ata ttt gat cgt gta ctc act gaa ctc gtt gct aaa atg 912
Val Glu Ala Ile Phe Asp Arg Val Leu Thr Glu Leu Val Ala Lys Met
290 295 300

aga gat atg ggt atg gat aga aca gag ctt ggc tgt ttg cgt act att 960
Arg Asp Met Gly Met Asp Arg Thr Glu Leu Gly Cys Leu Arg Thr Ile
305 310 315 320

att ctt ttt aat cca ggt tca aaa ggt ttg cag tct gtg aat gaa gtg 1008
Ile Leu Phe Asn Pro Gly Ser Lys Gly Leu Gln Ser Val Asn Glu Val
325 330 335

gaa gta ctg cgt gat aag gtt tat gtt gcg tta gaa gaa tat tgt cgt 1056
Glu Val Leu Arg Asp Lys Val Tyr Val Ala Leu Glu Glu Tyr Cys Arg
340 345 350

aca aca cat cca gaa gaa cct gga cga ttt gct aaa cta ctt ctt cgg 1104
Thr Thr His Pro Glu Glu Pro Gly Arg Phe Ala Lys Leu Leu Leu Arg
355 360 365

ctt cct tca tta cgt tca att gga tta aaa tgt ctg gaa cat tta ttc 1152
Leu Pro Ser Leu Arg Ser Ile Gly Leu Lys Cys Leu Glu His Leu Phe
370 375 380

ttt tat aaa ctt att ggc gat tcc cca att gat aca ttt tta atg gaa 1200
Phe Tyr Lys Leu Ile Gly Asp Ser Pro Ile Asp Thr Phe Leu Met Glu
385 390 395 400

gtt ctc gaa tca tct tca cat gac gtt caa gta gct aca tga 1242
Val Leu Glu Ser Ser Ser His Asp Val Gln Val Ala Thr
405 410

<210> 20
<211> 413
<212> PRT
<213> Myzus persicae

<400> 20
Met Asp Gly Thr Glu Arg Gly Leu Arg Leu Asp Asn Asn Leu Ser Leu
1 5 10 15

```
Ser Ser Met Gly Pro Gln Ser Pro Leu Asp Leu Lys Pro Asp Thr Ala
            20                  25                  30

Thr Leu Met Val Asn Phe Ser Pro Pro Gly Ala Pro Leu Ser Pro Ala
        35                  40                  45

Gly Leu Tyr Ser Val Asp Arg Asn Ser Met Met Asn Asn Ser Cys Asn
    50                  55                  60

Val Gln Asp Ser Pro Asn Tyr Pro Pro Asn His Pro Leu Ser Gly Ser
65                  70                  75                  80

Lys His Leu Cys Ser Ile Cys Gly Asp Arg Ala Ser Gly Lys His Tyr
                85                  90                  95

Gly Val Tyr Ser Cys Glu Gly Cys Lys Gly Phe Phe Lys Arg Thr Val
            100                 105                 110

Arg Lys Asn Leu Ser Tyr Ala Cys Arg Glu Glu Asn Lys Cys Ile Ile
        115                 120                 125

Asp Lys Arg Gln Arg Asn Arg Cys Gln Tyr Cys Arg Tyr Gln Lys Cys
    130                 135                 140

Leu Thr Met Gly Met Lys Arg Glu Ala Val Gln Glu Glu Arg Gln Arg
145                 150                 155                 160

Thr Lys Glu Arg Asp His Asn Asn Ile Glu Val Glu Pro Thr Ser Ser
                165                 170                 175

Ser Asn Thr Asp Met Pro Val Glu Leu Ile Leu Arg Ala Glu Asn Lys
            180                 185                 190

Ala Asp Ala Ile Lys Thr Glu Gln Gln Tyr Ile Glu Gln Arg His Pro
        195                 200                 205

Gln His Thr Val Gly Ala Ile Cys Gln Ala Thr Asp Lys Gln Leu Ile
    210                 215                 220

Gln Leu Val Glu Trp Ala Lys His Ile Pro His Phe Lys Asn Leu Pro
225                 230                 235                 240

Leu Gly Asp Gln Val Leu Leu Leu Arg Ala Gly Trp Asn Glu Leu Met
                245                 250                 255

Ile Ala Ala Phe Ser His Arg Ser Ile Ser Val Lys Asp Gly Ile Val
```

260 265 270

Leu Ala Thr Gly Leu Thr Val Asp Arg Asp Ser Ala His Gln Ala Gly
275 280 285

Val Glu Ala Ile Phe Asp Arg Val Leu Thr Glu Leu Val Ala Lys Met
290 295 300

Arg Asp Met Gly Met Asp Arg Thr Glu Leu Gly Cys Leu Arg Thr Ile
305 310 315 320

Ile Leu Phe Asn Pro Gly Ser Lys Gly Leu Gln Ser Val Asn Glu Val
325 330 335

Glu Val Leu Arg Asp Lys Val Tyr Val Ala Leu Glu Glu Tyr Cys Arg
340 345 350

Thr Thr His Pro Glu Glu Pro Gly Arg Phe Ala Lys Leu Leu Leu Arg
355 360 365

Leu Pro Ser Leu Arg Ser Ile Gly Leu Lys Cys Leu Glu His Leu Phe
370 375 380

Phe Tyr Lys Leu Ile Gly Asp Ser Pro Ile Asp Thr Phe Leu Met Glu
385 390 395 400

Val Leu Glu Ser Ser Ser His Asp Val Gln Val Ala Thr
405 410

<210> 21
<211> 150
<212> DNA
<213> Lucilia cuprina

<220>
<221> CDS
<222> (9)..(134)

<400> 21
aattctgc gaa gga tgc aag gga ttc ttc aaa cgt acc gta cgc aag gac 50
Glu Gly Cys Lys Gly Phe Phe Lys Arg Thr Val Arg Lys Asp
1 5 10

ttg aca tat gct tgt cgt gag gac aga aat tgc att ata gat aaa cga 98
Leu Thr Tyr Ala Cys Arg Glu Asp Arg Asn Cys Ile Ile Asp Lys Arg
15 20 25 30

caa aga aat cgt tgc cag tat tgt cgc tac caa aag tgatcgatac cgtcga 150
Gln Arg Asn Arg Cys Gln Tyr Cys Arg Tyr Gln Lys
35 40

<210> 22
<211> 42
<212> PRT
<213> Lucilia cuprina
<400> 22
Glu Gly Cys Lys Gly Phe Phe Lys Arg Thr Val Arg Lys Asp Leu Thr
1 5 10 15

Tyr Ala Cys Arg Glu Asp Arg Asn Cys Ile Ile Asp Lys Arg Gln Arg
20 25 30

Asn Arg Cys Gln Tyr Cys Arg Tyr Gln Lys
35 40

<210> 23
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:primer

<400> 23
cggaattccg cctcnggnta ycaytayaay gc 32

<210> 24
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:primer

<400> 24
cgcggatccr cactcctgac actttcgyct ca 32

<210> 25
<211> 23
<212> DNA

<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:primer

<400> 25
gcctcggggt atcactataa cgc 23

<210> 26
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:primer

<400> 26
gcactcctga cactttcgtc tca 23

<210> 27
<211> 23
<212> DNA

<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:primer

<400> 27
tcgtccggtt accattacaa cgc 23

<210> 28
<211> 24
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence:primer

<400> 28
tagacctttg gcraaytcna caat 24

<210> 29
<211> 37

<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:oligonucleotide

<400> 29
tcgacatata acttcgctgc agatgcatcc gagctct 37

<210> 30
<211> 35
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:oligonucleotide

<400> 30
ctagagctcg gatgcatctg cagcgaagtt atatg 35

<210> 31
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:oligonucleotide

<400> 31
tccagaaccg cggatagata tctgggatcc tc 32

<210> 32
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:oligonucleotide

<400> 32
ggagaggatc ccagatatct atccgcggtt ct 32

<210> 33

<211> 58
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:oligonucleotide

<400> 33
gatccatggg acaccatcac catcaccata ggccttccga acgcggtgaa ttccgaca 58

<210> 34
<211> 58
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:oligonucleotide

<400> 34
agcttgtcgg aattcaccgc gttcggaagg cctatggtga tggtgatggt gtcccatg 58

<210> 35
<211> 42
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:oligonucleotide

<400> 35
ccgggatctc gagatggact acaaggacga cgatgacaag cc 42

<210> 36
<211> 42
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:oligonucleotide

<400> 36
catgggcttg tcatcgtcgt ccttgtagtc catctcgaga tc 42

<210> 37
<211> 91
<212> DNA
<213> Bemisia tabaci

<220>
<221> CDS
<222> (2)..(91)

<400> 37
```
c aag agg aca gtt cgg aaa gac ttg tct tat gct tgc cgt gaa gaa aag 49
  Lys Arg Thr Val Arg Lys Asp Leu Ser Tyr Ala Cys Arg Glu Glu Lys
    1               5                  10                  15

aac tgt ctg att gat aag agg caa agg aat cga tgt caa tat            91
Asn Cys Leu Ile Asp Lys Arg Gln Arg Asn Arg Cys Gln Tyr
             20                  25                  30
```

<210> 38
<211> 30
<212> PRT
<213> Bemisia tabaci

<400> 38
```
Lys Arg Thr Val Arg Lys Asp Leu Ser Tyr Ala Cys Arg Glu Glu Lys
  1               5                  10                  15

Asn Cys Leu Ile Asp Lys Arg Gln Arg Asn Arg Cys Gln Tyr
             20                  25                  30
```

<210> 39
<211> 1491
<212> DNA
<213> Bemisia tabaci

<220>
<221> CDS
<222> (1)..(1488)

<400> 39
```
atg agt gaa aaa gaa aat gaa gca aaa aag att aaa ctt gat tcc tct   48
Met Ser Glu Lys Glu Asn Glu Ala Lys Lys Ile Lys Leu Asp Ser Ser
  1               5                  10                  15

ggt att ggt ata ata gaa tca tcc gaa tct aac gga gct atc atg ttg   96
```

```
Gly Ile Gly Ile Ile Glu Ser Ser Glu Ser Asn Gly Ala Ile Met Leu
            20                  25                  30

aac gat tct aca tct aat tgt cca tca cct tca cca tct cgt gtg gtt    144
Asn Asp Ser Thr Ser Asn Cys Pro Ser Pro Ser Pro Ser Arg Val Val
        35                  40                  45

cat ata cgg aat gtt ccc ata gaa gct act gaa aat gat gtt ctc agt    192
His Ile Arg Asn Val Pro Ile Glu Ala Thr Glu Asn Asp Val Leu Ser
    50                  55                  60

att ggt act cca ttt ggt gag atc acc aat gtt ctt ttg gtg aga gga    240
Ile Gly Thr Pro Phe Gly Glu Ile Thr Asn Val Leu Leu Val Arg Gly
65                  70                  75                  80

aaa ggt caa gcc ttc tta gag ttt gtt gac tca ttc tct gct caa caa    288
Lys Gly Gln Ala Phe Leu Glu Phe Val Asp Ser Phe Ser Ala Gln Gln
                85                  90                  95

atg gtt aac tgt tgg tct gat cct aac aac tca ccg atg caa ctt tgt    336
Met Val Asn Cys Trp Ser Asp Pro Asn Asn Ser Pro Met Gln Leu Cys
            100                 105                 110

att cgg gga aga caa gta tgt gtc cag ttt tca aag cac aaa gaa ctt    384
Ile Arg Gly Arg Gln Val Cys Val Gln Phe Ser Lys His Lys Glu Leu
        115                 120                 125

aaa aaa tct ctt ctt gga aca aat gct ggt tca gac agc agc tat caa    432
Lys Lys Ser Leu Leu Gly Thr Asn Ala Gly Ser Asp Ser Ser Tyr Gln
    130                 135                 140

agt aca tct cct caa aat agt aga cat ata agc aac ggt gat tct gtt    480
Ser Thr Ser Pro Gln Asn Ser Arg His Ile Ser Asn Gly Asp Ser Val
145                 150                 155                 160

gga gca agt tcc gtt ttc tct aat cca aat cat cct tta agc gga tca    528
Gly Ala Ser Ser Val Phe Ser Asn Pro Asn His Pro Leu Ser Gly Ser
                165                 170                 175

aaa cat ctc tgt tct att tgt ggt gat cga gcc tct ggg aaa cat tat    576
Lys His Leu Cys Ser Ile Cys Gly Asp Arg Ala Ser Gly Lys His Tyr
            180                 185                 190

ggt gtt tac agt tgt gaa gga tgt aaa gga ttt ttt aaa agg act gtt    624
Gly Val Tyr Ser Cys Glu Gly Cys Lys Gly Phe Phe Lys Arg Thr Val
        195                 200                 205
```

```
cgt aaa gat ttg tct tat gct tgt cgg gaa gaa cga gat tgt atc ata    672
Arg Lys Asp Leu Ser Tyr Ala Cys Arg Glu Glu Arg Asp Cys Ile Ile
    210                 215                 220

gac aga cga caa agg aat agg tgt caa tac tgt aga tat cag aaa tgt    720
Asp Arg Arg Gln Arg Asn Arg Cys Gln Tyr Cys Arg Tyr Gln Lys Cys
225                 230                 235                 240

ctc gct atg gga atg aaa aga gaa gcc gtg caa gaa gaa aga caa agg    768
Leu Ala Met Gly Met Lys Arg Glu Ala Val Gln Glu Glu Arg Gln Arg
                245                 250                 255

aat aaa gaa aaa agt gaa aac gag gtt gaa agt aca agt aac tca cag    816
Asn Lys Glu Lys Ser Glu Asn Glu Val Glu Ser Thr Ser Asn Ser Gln
            260                 265                 270

aat gat atg cct atc gaa aga ata ctg gaa gct gaa tta cga gtg gaa    864
Asn Asp Met Pro Ile Glu Arg Ile Leu Glu Ala Glu Leu Arg Val Glu
        275                 280                 285

cct aag aat gaa gac ata gat tct cga gat ccc gtt agt gat atc tgt    912
Pro Lys Asn Glu Asp Ile Asp Ser Arg Asp Pro Val Ser Asp Ile Cys
    290                 295                 300

caa gcg gca gat cga caa ctt tac caa tta att gaa tgg gct aag cat    960
Gln Ala Ala Asp Arg Gln Leu Tyr Gln Leu Ile Glu Trp Ala Lys His
305                 310                 315                 320

att cct cat ttc acc gag tta ccc gtt gaa gat caa gtt att tta ctt    1008
Ile Pro His Phe Thr Glu Leu Pro Val Glu Asp Gln Val Ile Leu Leu
                325                 330                 335

aaa tca gga tgg aat gag ctt ctc att gca ggc ttt tct cat cgt tca    1056
Lys Ser Gly Trp Asn Glu Leu Leu Ile Ala Gly Phe Ser His Arg Ser
            340                 345                 350

atg tca gtt aaa gat ggt atc atg tta gcc act ggt ttg gtt gtt cat    1104
Met Ser Val Lys Asp Gly Ile Met Leu Ala Thr Gly Leu Val Val His
        355                 360                 365

aga aac tgt gct cat caa gcg ggt gta ggt gct att ttt gat cgc gtg    1152
Arg Asn Cys Ala His Gln Ala Gly Val Gly Ala Ile Phe Asp Arg Val
    370                 375                 380

tta act gaa tta gtg gct aaa atg aga gaa atg aaa atg gac aaa act    1200
```

```
Leu Thr Glu Leu Val Ala Lys Met Arg Glu Met Lys Met Asp Lys Thr
385                 390                 395                 400

gaa ctt ggt tgc tta cga tct att gta tta ttt aat ccc gaa gct aaa     1248
Glu Leu Gly Cys Leu Arg Ser Ile Val Leu Phe Asn Pro Glu Ala Lys
                405                 410                 415

gga ctc aaa tca aca caa caa gtt gaa aat tta cgt gaa aag gtt tac     1296
Gly Leu Lys Ser Thr Gln Gln Val Glu Asn Leu Arg Glu Lys Val Tyr
            420                 425                 430

gca atc ctg gaa gag tat tgt aga caa act tat cct gat caa tct ggc     1344
Ala Ile Leu Glu Glu Tyr Cys Arg Gln Thr Tyr Pro Asp Gln Ser Gly
        435                 440                 445

cgt ttt gct aaa tta ctt ctt cgt cta cct gcc tta cgg tca att ggt     1392
Arg Phe Ala Lys Leu Leu Leu Arg Leu Pro Ala Leu Arg Ser Ile Gly
    450                 455                 460

tta aag tgt ttg gaa cat tta ttt ttc ttc aaa ttg gtt gga aac aca     1440
Leu Lys Cys Leu Glu His Leu Phe Phe Phe Lys Leu Val Gly Asn Thr
465                 470                 475                 480

tct att gac agt ttc ttg ttg tcc atg tta gaa tct aat tca gac tca     1488
Ser Ile Asp Ser Phe Leu Leu Ser Met Leu Glu Ser Asn Ser Asp Ser
                485                 490                 495

tag                                                                 1491


<210> 40
<211> 496
<212> PRT
<213> Bemisia tabaci

<400> 40
Met Ser Glu Lys Glu Asn Glu Ala Lys Lys Ile Lys Leu Asp Ser Ser
  1               5                  10                  15

Gly Ile Gly Ile Ile Glu Ser Ser Glu Ser Asn Gly Ala Ile Met Leu
             20                  25                  30

Asn Asp Ser Thr Ser Asn Cys Pro Ser Pro Ser Pro Ser Arg Val Val
        35                  40                  45

His Ile Arg Asn Val Pro Ile Glu Ala Thr Glu Asn Asp Val Leu Ser
    50                  55                  60
```

```
Ile Gly Thr Pro Phe Gly Glu Ile Thr Asn Val Leu Leu Val Arg Gly
 65                  70                  75                  80

Lys Gly Gln Ala Phe Leu Glu Phe Val Asp Ser Phe Ser Ala Gln Gln
                 85                  90                  95

Met Val Asn Cys Trp Ser Asp Pro Asn Asn Ser Pro Met Gln Leu Cys
            100                 105                 110

Ile Arg Gly Arg Gln Val Cys Val Gln Phe Ser Lys His Lys Glu Leu
        115                 120                 125

Lys Lys Ser Leu Leu Gly Thr Asn Ala Gly Ser Asp Ser Ser Tyr Gln
    130                 135                 140

Ser Thr Ser Pro Gln Asn Ser Arg His Ile Ser Asn Gly Asp Ser Val
145                 150                 155                 160

Gly Ala Ser Ser Val Phe Ser Asn Pro Asn His Pro Leu Ser Gly Ser
                165                 170                 175

Lys His Leu Cys Ser Ile Cys Gly Asp Arg Ala Ser Gly Lys His Tyr
            180                 185                 190

Gly Val Tyr Ser Cys Glu Gly Cys Lys Gly Phe Phe Lys Arg Thr Val
        195                 200                 205

Arg Lys Asp Leu Ser Tyr Ala Cys Arg Glu Glu Arg Asp Cys Ile Ile
    210                 215                 220

Asp Arg Arg Gln Arg Asn Arg Cys Gln Tyr Cys Arg Tyr Gln Lys Cys
225                 230                 235                 240

Leu Ala Met Gly Met Lys Arg Glu Ala Val Gln Glu Glu Arg Gln Arg
                245                 250                 255

Asn Lys Glu Lys Ser Glu Asn Glu Val Glu Ser Thr Ser Asn Ser Gln
            260                 265                 270

Asn Asp Met Pro Ile Glu Arg Ile Leu Glu Ala Glu Leu Arg Val Glu
        275                 280                 285

Pro Lys Asn Glu Asp Ile Asp Ser Arg Asp Pro Val Ser Asp Ile Cys
    290                 295                 300
```

```
Gln Ala Ala Asp Arg Gln Leu Tyr Gln Leu Ile Glu Trp Ala Lys His
305                 310                 315                 320

Ile Pro His Phe Thr Glu Leu Pro Val Glu Asp Gln Val Ile Leu Leu
                325                 330                 335

Lys Ser Gly Trp Asn Glu Leu Leu Ile Ala Gly Phe Ser His Arg Ser
            340                 345                 350

Met Ser Val Lys Asp Gly Ile Met Leu Ala Thr Gly Leu Val Val His
        355                 360                 365

Arg Asn Cys Ala His Gln Ala Gly Val Gly Ala Ile Phe Asp Arg Val
    370                 375                 380

Leu Thr Glu Leu Val Ala Lys Met Arg Glu Met Lys Met Asp Lys Thr
385                 390                 395                 400

Glu Leu Gly Cys Leu Arg Ser Ile Val Leu Phe Asn Pro Glu Ala Lys
                405                 410                 415

Gly Leu Lys Ser Thr Gln Gln Val Glu Asn Leu Arg Glu Lys Val Tyr
            420                 425                 430

Ala Ile Leu Glu Glu Tyr Cys Arg Gln Thr Tyr Pro Asp Gln Ser Gly
        435                 440                 445

Arg Phe Ala Lys Leu Leu Leu Arg Leu Pro Ala Leu Arg Ser Ile Gly
    450                 455                 460

Leu Lys Cys Leu Glu His Leu Phe Phe Phe Lys Leu Val Gly Asn Thr
465                 470                 475                 480

Ser Ile Asp Ser Phe Leu Leu Ser Met Leu Glu Ser Asn Ser Asp Ser
                485                 490                 495


<210> 41
<211> 101
<212> DNA
<213> Bemisia tabaci

<220>
<221> CDS
<222> (2)..(100)
```

<400> 41
c ctc acc tgc gaa ggc tgc aag ggc ttc ttc cgt cgg agc atc acc aag 49
Leu Thr Cys Glu Gly Cys Lys Gly Phe Phe Arg Arg Ser Ile Thr Lys
1 5 10 15

aat gcc gtc tac cag tgt aaa tat gga aat aat tgt gaa atc gac atg 97
Asn Ala Val Tyr Gln Cys Lys Tyr Gly Asn Asn Cys Glu Ile Asp Met
20 25 30

tac a 101
Tyr

<210> 42
<211> 33
<212> PRT
<213> Bemisia tabaci

<400> 42
Leu Thr Cys Glu Gly Cys Lys Gly Phe Phe Arg Arg Ser Ile Thr Lys
1 5 10 15

Asn Ala Val Tyr Gln Cys Lys Tyr Gly Asn Asn Cys Glu Ile Asp Met
20 25 30

Tyr

INTERNATIONAL SEARCH REPORT

International application No.
PCT/AU00/00799

A. CLASSIFICATION OF SUBJECT MATTER

Int. Cl. ^{7}: C07K 14/72, C12N 15/12, C07H 21/04

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
AS ABOVE

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
AS BELOW

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
Gen-Pept; SWISS-PROT; PIR; TREMBL : Sequence ID Nos. 2, 4, 6, 8, 10, 14, 16, 18, 20, 40

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	SWISS-PROT Accession no. 018531, 15 December 1998, Hannan G. N., Hill R. J.	1-5,9,10,15-21,27-40,44 Seq. ID 2
X	SWISS-PROT Accession no. P34021, 1 February 1994, Koelle M. R., *et.al.*	1-5,9,10,15-21,27-40,44 Seq.ID 2
X	GenPept Accession no. CAA36827, 1990, Henrich, V. C.	1-5,9,11,15-22,27-40,44 Seq ID 4,6,8

[X] Further documents are listed in the continuation of Box C [X] See patent family annex

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search 8 August 2000	Date of mailing of the international search report 29 AUG 2000
Name and mailing address of the ISA/AU AUSTRALIAN PATENT OFFICE PO BOX 200, WODEN ACT 2606, AUSTRALIA E-mail address: pct@ipaustralia.gov.au Facsimile No. (02) 6285 3929	Authorized officer **ALISTAIR BESTOW** Telephone No : (02) 6283 2450

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	GenPept Accession no. CAA72296, 1997, Mouillet J. F. *et.al.*	1-7,19,20, 23, 27-40,44 Seq ID 10,14
X	SWISS-PROT Accession no. P20153; Q9W535, 1 February 1991, Oro A. E., *et.al.*	1-5,9,11,15-22,27-40,44 Seq ID 4,6,8
X	GenPept Accession no. AAD19828, 1998, Saleh D. S. *et.al.*	1-7,19,20, 23, 27-40,44 Seq ID 10,14
X	GenPept Accession no. AAB94565, 1997, Guo X., *et.al.*	1-7,19,20, 23, 27-40,44 Seq ID 10
X	GenPept Accession no. AAB94566, 1997, Guo X., *et.al*	1-7,19,20, 23, 27-40,44 Seq ID 10,14
X	GenPept Accession no. AAC95154, 13 October 1998, Vaillancourt, P. *et.al.*	1-7,19,20, 23,27-40,44 Seq ID 16,18,20
X	GenPept Accession no. AAC59722, 1995, Jones B. B. *et.al.*	1,2,15-20,27-40,44 Seq ID 40
X	SWISS-PROT Accession no P28700, 1 December 1992, Leid M. *et.al.*	1-7,19,20, 23,27-40,44 Seq ID 16,18,20,40
X	SWISS-PROT Accession no P19793, 1 February 1991, Manglesdorf D.J., *et. al.*	1-7,19,20, 23,27-40,44 Seq ID 16,18,20,40
X	GenPept Accession no AAC59721, 1995, Jones B. B., *et.al.*	1-7,19,20, 23,27-40,44 Seq ID 20,40
P,X	WO, A, 99/36520 (COMMONWEALTH AND SCIENTIFIC AND INDUSTRIAL RESEARCH ORGANISATION.) 22 July 1999	1-5,9,11,15-22,27-40,44 Seq. 2,4,6,8,10

C (Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO, A, 98/35550 (NEW ZEALAND PASTORAL AGRICULTURE RESEARCH INSTITUTE LIMITED and THE NEW ZEALAND WOOL BOARD) 20 August 1998	1-5,9,11,15-22,27-40,44 Seq ID 2,4,6
P,X	US, A, 6025483 (THE REGENTS OF THE UNIVERSITY OF CALIFORNIA) 15 February 2000	1-5,9,11,15-22,27-40,44 Seq ID 6
X	WO, A, 91/13167 (THE BOARD OF TRUSTEES OF LELAND STANFORD Jnr. UNIVERSITY) 18 September 1991	1-5,9,11,15-22,27-40,44 Seq ID 4
P,X	WO, A, 99/48915 (GLAXO GROUP LIMITED) 18 October 1999	1-7,19,20,23,27-40,44 Seq ID 16

Box I **Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)**

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos :

because they relate to subject matter not required to be searched by this Authority, namely:

2. ☒ Claims Nos : 41 - 43

because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically: The method claims are not restricted to the use of the steroid receptors of the invention. The compound claim is characterised in terms of a structure governed by the interaction with a receptor or a fragment thereof, such a receptor not being restricted to that of the invention. This is unsearchable.

3. ☐ Claims Nos :

because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a)

Box II **Observations where unity of invention is lacking (Continuation of item 3 of first sheet)**

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest ☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/AU00/00799

This Annex lists the known "A" publication level patent family members relating to the patent documents cited in the above-mentioned international search report. The Australian Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

		Patent Family Member					
WO	9948915	AU	32116/99				
WO	9113167	AU	74922/91	AU	17792/95	AU	49218/97
		CA	2076386	EP	517805	US	5514578
US	6025483	NONE					
WO	9835550	AU	60077/98				
WO	9936520	AU	21429/99				

END OF ANNEX